Force-spectroscopy of small ligands binding to nucleic acids

Joan Camuñas i Soler

JOAN CAMUÑAS I SOLER

FORCE-SPECTROSCOPY OF SMALL LIGANDS BINDING TO NUCLEIC ACIDS

FORCE-SPECTROSCOPY OF SMALL LIGANDS BINDING TO NUCLEIC ACIDS

JOAN CAMUÑAS I SOLER

DEPARTAMENT DE FÍSICA FONAMENTAL
FACULTAT DE FÍSICA
UNIVERSITAT DE BARCELONA

PROGRAMA DE DOCTORAT EN FÍSICA

Memòria presentada el 19 de desembre de 2014 per optar al títol de

DOCTOR EN FÍSICA

sota la direcció i tutorització del

DR. FÈLIX RITORT I FARRAN

TESI DIRIGADA PER:
Dr. Fèlix Ritort i Farran

SUMMARY

Single-molecule techniques allow to follow biomolecular reactions with unprecedented resolution. Particularly, optical tweezers allow to manipulate and apply forces to individual molecules tethered between plastic or silica beads by using laser beams. Typical experiments consist in manipulating nucleic acids (DNA, RNA) or proteins. For instance DNA molecules can be stretched to measure its elastic properties, or unzipped to measure their base-pairing energies. Many anticancer drugs target nucleic acids to exert their cytotoxic activity. To understand their mechanism of action it is important to know at which positions, how strong, and how fast they bind different sites in DNA. Single-molecule optical tweezers experiments can be used to unravel the binding thermodynamics and kinetics of these ligands, specially those difficult to characterize with bulk techniques. Thiocoraline is one of such drugs, and binds DNA through bis-intercalation. Experiments with optical tweezers show that the kinetics of intercalation are very slow (hours) and strongly force-dependent: force facilitates binding but slows down unbinding. Experiments also reveal that the binding pathway proceeds through a mono-intercalated intermediate responsible of the slow kinetics. DNA unzipping is used to determine the preferred binding sequences of Thiocoraline finding that it preferentially clamps CpG steps. This represents a single-molecule footprinting technique that can be parallelized using magnetic tweezers. The fact that bis-intercalation does not modify DNA persistence length is also found from pulling experiments. This combination of DNA stretching and unzipping assays is also used to follow how the anticancer agent Kahalalide F self-assembles and compacts DNA. Kahalalide F forms positively-charged nanometric particles able to bind and condense DNA. The binding reaction shows to phases: an initial compaction of electrostatic origin, and a subsequent stiffening due to the hydrophobic collapse of the complex. Modeling of the experiments provides the thermodynamic parameters of the interaction that are complemented with kinetic measurements. A simple methodology to produce ssDNA with optical tweezers is used to study how the stiffness of the polyanion affects its compaction, and then to study the elasticity of ssDNA under varying ionic conditions. The utility of these techniques to study self-assembly and aggregation is finally explored with amyloidogenic peptides involved in neurodegenerative disorders.

CONTENTS

Part I

INTRODUCTION AND OBJECTIVES

> Watching a coast as it slips by the ship is like thinking about an enigma. There it is before you, smiling, frowning, inviting, grand, mean, insipid, or savage, and always mute with an air of whispering, "Come and find out".
>
> *Joseph Conrad (Heart of Darkness)*

CHAPTER 1

SINGLE-MOLECULE BIOPHYSICS

Most of the accumulated knowledge on the functioning and structure of biomolecules over the XXth century has been obtained from biochemical bulk assays in which the ensemble properties of a large number of molecules are measured simultaneously. A typical bulk experiment might involve following a chemical reaction by measuring the changes in fluorescence of a small volume ($\sim$ µl) of a molecular sample at µM concentration. Even in such conditions more than 10^{12} molecules are contained in the sample and contribute to the measured signal. Although much can be learnt from the population averages obtained in such classical ensemble methods, also some important features are lost. An example where such features hidden to macroscopic observations are particularly relevant is the translocation kinetics of molecular motors (Figure 1). For instance, the dynamics of an RNA polymerase transcribing DNA in a bulk assay might look like a smooth and continuously varying polymerization reaction due to the unsynchronized average over a large number of transcribing enzymes (Figure 1**a**). Although such experiments provide the average transcription speed of the enzyme, this picture is far from the underlying molecular dynamics of an RNA polymerase.

Single-molecule techniques allow us to follow these enzymes individually giving access to their translocation kinetics, and to the coupling between the observed mechanical steps and the chemical reactions that drive them (i.e. the mechano-chemical cycle). Indeed single-molecule experiments have shown that the motion of polymerases is stochastic, showing pauses and changes in its velocity while translo-

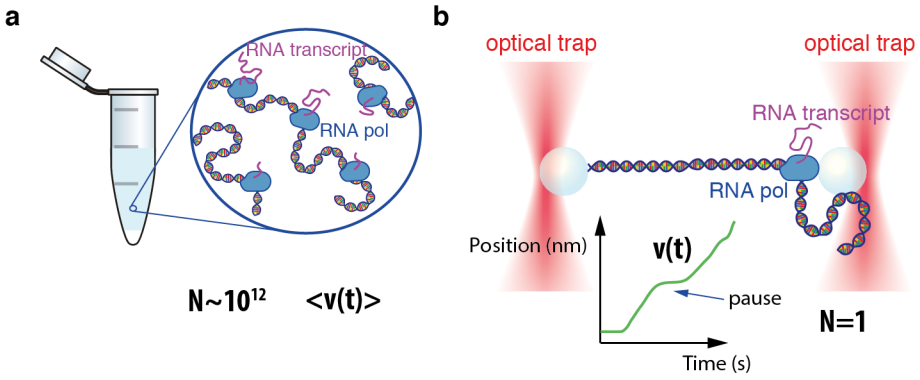

Figure 1: Transcription as observed in bulk and single molecule. (**a**) In a bulk experiment, measurements are simultaneously performed over a large number of enzymes, N, each of them in a different stage of the transcription reaction. Average properties such as the transcription speed $\langle v(t) \rangle$ can be determined. (**b**) In single-molecule experiments, a single RNA polymerase is followed while translocating along DNA. In this way, the transcription speed, v(t), at different regions (e.g CG-rich) and stages (e.g. initiation, termination) can be followed in real-time. Particularly, using optical tweezers, an opposing force, F, can be applied to the enzyme, giving access to features unobservable in bulk setups such as the stalling force of the enzyme.

cating along a DNA template (Figure 1**b**). Such information would be very difficult to obtain without a single-molecule approach. Moreover, most biomolecular reactions take place in the energy scale of thermal fluctuations (a few k_BT), and therefore these fluctuations are sometimes an essential component of their function. In fact, individual molecules in a thermal bath usually explore conformational states that are far from their macroscopic average values. When such kinetic states are important to understand a molecular process, a microscopic view is needed, and single-molecule tools become useful. Similarly, the detection of molecular states or conformations that are underrepresented in a population (e.g. intermediate states in a protein folding reaction) can be challenging with bulk techniques, as their signal becomes masked by those of the most abundant or longer-lived species. Consequently, the development of techniques to observe and manipulate molecules one at a time have allowed to observe molecular heterogeneity at an unprecedented level, challenging previous knowledge in molecular biophysics. Moreover, with these single-molecule techniques it is possible to measure physical magnitudes difficult to determine with previous techniques (such as force and torque), giving access to the elementary biochemical steps of a large number of experimental systems.

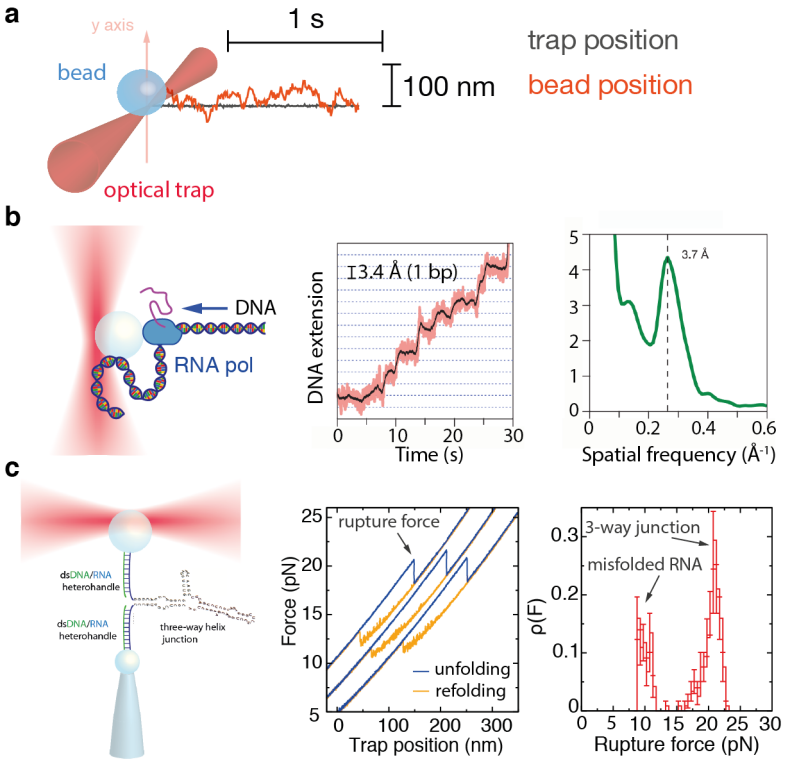

Figure 2: Fluctuations in single-molecule experiments. (**a**) A bead in an optical trap fluctuates around its equilibrium position (bead not to scale). For a trap with very low stiffness (k ~ 4 pN/μm), fluctuations can be as large as 100 nm. (**b**) Following with the RNA polymerase example, high frequency data on the translocation dynamics of the enzyme (middle panel, red) can be filtered down (black) in order to increase the signal-to noise ratio. For a process that is slow enough, this allows to detect subnanometric steps. A Fourier transformation (right panel) shows that the enzyme moves in steps of 1 bp (3.4 Å). Adapted from [1]. (**c**) An RNA three-way junction can be pulled with optical tweezers. In pulling experiments (middle) the force required to unfold the structure is stochastic. The distribution of rupture forces over a large number of realizations (right) shows a bimodal distribution indicative of a misfolded structure that competes with the native one.

1.1 THE PHYSICS OF SMALL SYSTEMS

A fundamental feature of single-molecule experiments is the observation of fluctuations in the measured quantities. These fluctuations rarely play a role in typical macroscopic systems containing a large number of molecules or subunits N. Energy fluctuations scale as $1/\sqrt{N}$ and therefore, for macroscopic systems that exchange large energies with the bath, fluctuations are negligible. For instance, in the previous example with $N \sim 10^{12}$ molecules, measurements deviate only a very small fraction from the average ($\sim 10^{-6}$), which is undetectable in a typical spectrophotometric experiment. However, when only a few molecules are being observed, fluctuations are large enough to be measured in the experimental system. Typical single-molecule experiments deal with systems (e.g. molecular motors, poly-

mer elasticity) that exchange energies on the order of 1-1000 k_BT with the bath, and consequently fluctuations are directly observed in the measured quantities [2]. This is due to the fact that the typical length scale of most biomolecular systems are in the range of nanometers, whereas forces are on the pN range. Hence, the energies that these reactions involve are on the order of a few k_BT's ($k_BT = 4.11$ pN·nm). These fluctuations provide important information from the molecular system under study. In this sense, the development during the 1990's of fluctuation theorems has provided valuable tools to obtain thermodynamic quantities from fluctuating non equilibrium measurements obtained in single-molecule experiments [2, 3].

In fact, in force-spectroscopy setups the probe itself is a small system subjected to thermal fluctuations (Figure 2**a**). Consequently, fluctuations observed in a single molecule experiment provide both information of the molecular system under study and of the force-probe. Uncoupling these contributions is an important part in many single-molecule techniques. For instance optical tweezers can be calibrated by measuring the spectrum of fluctuations of a trapped particle (see Appendix A.2.1). The observation of fluctuations also determines how single-molecule data is obtained and analyzed: whereas bulk measurements are the result of a simultaneous average over a large number of particles N, typical single-molecule measurements rely on two types of averages. In some experiments the trajectories described by individual molecules are integrated over time to increase the signal-to-noise ratio (e.g detecting Å-scale steps in molecular motors, Figure 2**b**). In other experiments in which information is extracted from the fluctuating quantities, the same experiment is repeated several times to measure distributions of a molecular property (e.g. rupture force distribution of a nucleic acid structure, Figure 2**c**).

1.2 EXPERIMENTAL TECHNIQUES

1.2.1 *Force-sensing techniques: atomic force microscopy, optical and magnetic tweezers*

Among the single-molecule micromanipulation techniques developed in the past years, three of them stand out due to their widespread use and versatility: atomic force microscopy (AFM), optical tweezers and magnetic tweezers. These techniques allow to manipulate single molecules and can be used as force-transducers to measure and apply forces to the tethered molecules. All together they cover the whole range of forces relevant to biomolecular reactions: from polymer entropic forces (several fN) up to the forces needed to break covalent bonds (a few nN). Here we will briefly discuss the main features of each technique (strengths and limitations) and their typical applications (summarized in Table 1).

Table 1: Comparison of single-molecule force-spectroscopy techniques [4].

Technique	AFM	Optical tweezers	Magnetic tweezers
Force Range (pN)	$10 - 10^4$	$0.1 - 100$	$0.01 - 100$
Spatial resolution (nm)	$0.5 - 1$	$0.1 - 2$	$5 - 10$
Stiffness (pN/nm)	$10 - 10^5$	$0.01 - 1$	$10^{-3} - 10^{-6}$
Temporal resolution (s)	10^{-3}	10^{-4}	$10^{-1} - 10^{-2}$
Probe size (μm)	$100 - 250$	$0.25 - 5$	$0.5 - 5$
Features	High-force pulling	Dumbbell geometry	Force clamp
	High-res. imaging	Versatile range	Bead rotation

1.2.1.1 *Atomic Force Microscopy (AFM)*

From the three force-spectroscopy techniques explained here, the AFM is arguably the most extended one due to its utility to perform images of non-conductive surfaces with Å-resolution. The AFM is based on the scanning tunneling microscope [5], and consists of a cantilever with a sharp tip at its end that is used to probe the surface under study. The cantilever bends when the tip is brought close to a surface due to the forces that arise between tip and sample. This deflection of the cantilever can be measured by reflecting a laser beam on its surface, and measuring the laser displacements with a position sensing detector (PSD) (Figure 3**a**). Typically, the sample is hold in a support stage that can be displaced with piezoelectric actuators. In this way, the roughness of a sample can be mapped with subnanometer resolution by measuring the deflection of the laser beam while the tip is displaced over the sample. Similarly, the height of the cantilever can also be adjusted with a piezoelectric actuator allowing to use feedback loops. Typical resolutions are on the scale of a few Å for the height profile (z-axis), and of $\sim$ 30 nm in the lateral plane (due to the convolution with the end radius of the tip).

AFM instruments can be operated in various modes: contact mode, tapping mode and jumping mode. In contact mode the deflection of the cantilever is kept constant using feedback while the sample is explored. This was the first scanning mode developed for AFMs and provides the highest spatial resolutions [6], but has an important disadvantage: scanning causes a frictional force that can damage soft samples such as biomaterials. To overcome this problem, the dynamic or "tapping" mode was developed. In this mode, the cantilever is vertically oscillated close to its natural resonance frequency, and approached to the surface (Figure 3**a**). When the tip of the cantilever is a few nanometers away from the surface, the amplitude of the oscillations sharply decreases. The surface can then be scanned while maintaining the amplitude of the oscillation constant using feedback. If the tip passes over a gap, the amplitude increases, and the feedback moves the tip towards the surface to keep the amplitude constant. Similarly, if the tip finds a protuberance, the amplitude decreases and

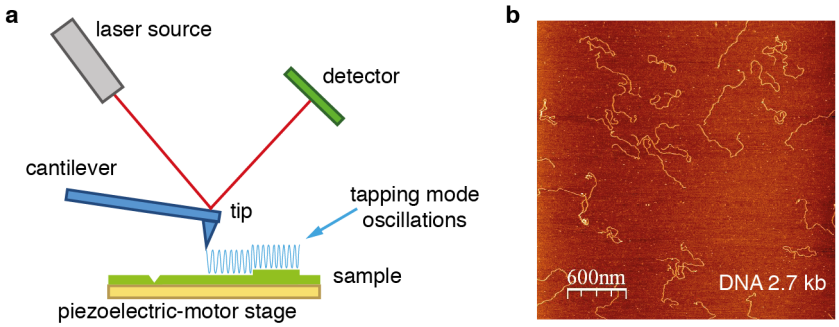

Figure 3: AFM imaging. (**a**) AFM configuration to image the roughness of a surface. The height of the sample (green) can be mapped using a cantilever with a sharp tip at its end (blue). Deflections are measured by reflecting a laser beam in the back of the cantilever. In the tapping mode, the tip is oscillated to minimize sample damage. (**b**) AFM image of DNA molecules (2.7 kb) adsorbed on a mica surface. DNA has a diameter of ~2 nm.

the feedback moves the tip away from the surface. In this way, the surface can be scanned minimizing both tip and sample damage [6].

Biological samples are usually prepared by deposition in a freshly cleaved mica surface using well established protocols (Figure 3**b**) [6]. The main advantages of AFM are therefore, the simple and rapid preparation of samples, and the capability of obtaining high-resolution images of biological samples in near-physiological liquid conditions [7]. Moreover, the recent development of high-speed AFM also allows to visualize dynamic processes of biomolecules in real time [8, 9]. Although AFM is mostly used as an imaging tool, it can also be used as a force-scanning probe to test molecular interactions in the force range between pN to nN, relevant to study the unfolding of large protein domains or antigen-antibody interactions (i.e. strong non-covalent interactions) [2, 7].

To perform single-molecule force measurements, the surface is usually coated with the molecules of interest (e.g. polyproteins, nucleic acids) using specific chemical attachments or non-specific adsorption. To pull on a molecule, the AFM tip is approached to the surface until a molecule is tethered between tip and surface. The tip is then retracted with the piezoelectric actuator at a constant pulling rate, and a force-extension curve recorded by measuring the deflection of the laser beam as a function of the tip displacement.

A major concern in AFM force-spectroscopy experiments is the non-specific interactions that can arise when the relatively large tip of the AFM is approached to the surface to pull on a single molecule. The use of non-specific attachments between molecule and tip are more straightforward to implement but usually lead to multiple tethers and larger uncertainties on the binding geometry of the experiment, complicating data interpretation. The use of specific chemical bonds (e.g. biotin-streptavidin, crosslinkers) and molecular spacers (called han-

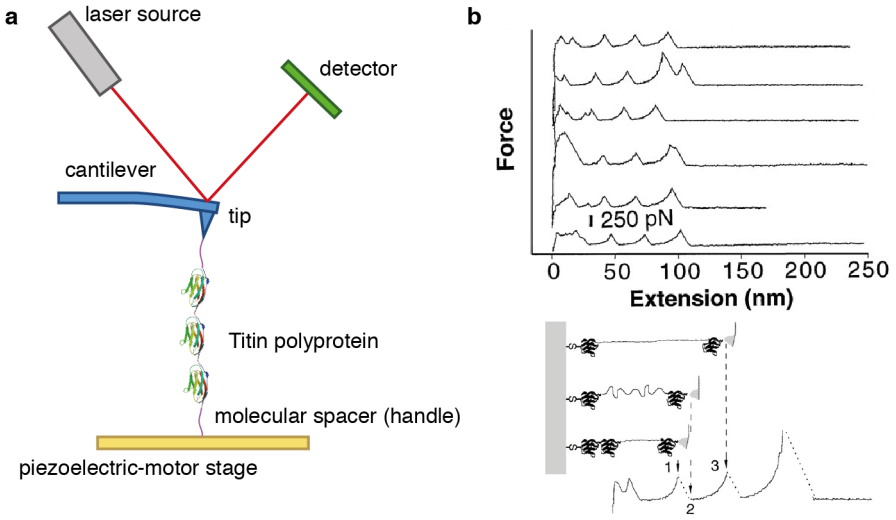

Figure 4: Force spectroscopy using AFM. (**a**) AFM setup in a protein pulling experiment. A polyprotein is tethered between the tip of the AFM and the surface using spacer handles. The molecule can be pulled by retracting the AFM tip, and the force is measured from the deflection of the cantilever using a laser beam. (**b**) Force-extension curves of the polyprotein (top) show a series of force peaks that correspond to the sequential unfolding of protein domains. In a typical pulling, the force force is increased up to 150-300 pN were unfolding of a protein domain becomes highly probable (bottom, 1). When a protein unfolds the force drops due to the released polypeptide chain (bottom, 2). Continued retraction of the tip stretches the released polypeptide chain until the next protein domain unfolds (bottom, 3). Adapted from [10].

dles) are more complex but reduce non-specific adhesions between tip and surface. To facilitate data interpretation, many times such linkers are designed to have a particular force fingerprint that allows to identify that a single molecule is being pulled (e.g. the opening of repeating units of a characteristic protein domain of Titin [10]).

Another intrinsic drawback of AFM force measurements arises from the relatively large size (~100 μm) and stiffness of cantilevers, which limits force resolution [4, 11]. A typical value for the stiffness of a cantilever is $k \sim 100$ pN/nm, and therefore from the fluctuation-dissipation relation it follows that force-fluctuations are on the order of $\sqrt{\langle \delta F^2 \rangle} = \sqrt{k_B T k} \sim 20$ pN (while distance fluctuations are on the Å-scale, $\sqrt{\langle \delta x^2 \rangle} = \sqrt{k_B T / k} \sim 2$ Å) [2]. Therefore, for a cantilever where noise is limited by thermal fluctuations, the signal-to-noise ratio ($SNR = \Delta F / \sqrt{\langle \delta F^2 \rangle}$, where ΔF is a change in force) is poor on the low pN range, that is where most weak molecular interactions take place (e.g DNA base-pair disruption).

For processes that vary slowly with time however, the measured signal can be filtered to lower frequencies increasing force resolution. For a probe working in overdamped conditions (e.g. soft AFM cantilevers in water and optical traps), the spectrum of fluctuations can

be derived from the equation of motion of an overdamped harmonic oscillator, finding:

$$\left\langle \delta x^2 \right\rangle = \int_0^{2\pi B} \frac{2k_B T}{\gamma \left(w^2 + w_c^2\right)} dw \tag{1}$$

where B is the bandwidth, γ is the viscous damping of the probe, and w_c is its corner frequency ($w_c = k/\gamma$). For filtering bandwidths below the corner frequency of the probe ($B < w_c$), the integrated mean square fluctuations are well approximated by $\left\langle \delta F^2 \right\rangle = k^2 \left\langle \delta x^2 \right\rangle \approx 4k_B TBk/w_c$. Consequently, a general limit to force resolution is set by a SNR $\geqslant 1$ in:

$$SNR = \frac{\Delta F}{\sqrt{4k_B TB\gamma}} \tag{2}$$

This is generally valid for molecular constructs softer than the transducer (as it is usually the case for AFMs where $k_{cantilever} >> k_{tether}$), and for setups in which the instrumental noise is smaller than thermal fluctuations. Consequently, for a given bandwidth, smaller cantilevers with lower damping coefficients γ provide better force resolution [11, 12]. Consequently, the recent development of ultrastable AFMs with smaller but less stiff tips, and active drift correction systems is allowing to overcome some of these limitations [12, 13]. However, so far force-spectroscopy with AFMs is still mostly used to investigate strong inter and intra molecular interactions (30 pN-10 nN) such as protein unfolding [10, 14], single-bond catalysis [15] and covalent bond disruption [16].

1.2.1.2 *Optical tweezers*

Optical tweezers are based in the principle that light carries a momentum that can be used to exert forces to microscopic dielectric particles. An optical trap is usually created by tightly focusing a laser beam with a high numerical aperture (NA) objective into a microfluidics chamber. Dielectric particles near the focused beam experience a net gradient force towards the focus, where they become confined. By displacing the focused beam, the optically-trapped particle can then be moved in the three-dimensions of space. To perform single-molecule manipulation experiments, usually polystyrene µm-sized beads are used for trapping, and the molecular construct of interest (e.g DNA, RNA, protein) is specifically attached between two optically-trapped beads (in dual-trap setups, Figure 1**b**) [17]. In more common single-trap setups however, one of the beads is subjected by air suction with a glass micropipette (Figure 2**c**), or alternatively a functionalized glass coverslip is used for one of the attachments.

Analogously to AFM cantilevers, an optical trap can also be used as a force-transducer to measure and apply forces to the trapped object. For small displacements of the particle from its equilibrium position

the trapping potential is harmonic, and therefore forces can be measured by considering the trap as a Hookean spring of stiffness k_{trap}, and measuring the displacement of the bead ($F = k_{trap}\Delta x$). The typical stiffness of optical traps is 1000-fold lower than that of AFM cantilevers leading to much lower force fluctuations. This allows to measure sub-pN force variations with high temporal resolution [2, 18]. Optical tweezers instruments can typically exert forces up to 100 pN, with very high force (sub-pN), spatial (sub-nanometer) and temporal (sub-millisecond) resolution, being probably the most versatile force-spectroscopy tool to test weak biomolecular interactions (e.g protein-DNA interactions, molecular motors, folding intermediates, base-pair disruption) (Figure 2). The natural control parameter of optical traps is the position of the trap, but force clamp measurements can also be performed using active feedback [19] and particular dual-trap configurations [20].

The main drawbacks of optical tweezers being the higher complexity to setup and align, and the difficulty to parallelize measurements or exert rotational control on the trapped bead. Nevertheless, innovative approaches are succeeding to solve some of these issues [21, 22]. As this thesis focuses in the use of optical tweezers to investigate peptide and protein-DNA interactions, a more detailed explanation of the principles, force-measurement methods and calibration of optical tweezers is found in Chapter 2.

1.2.1.3 *Magnetic tweezers*

Magnetic tweezers are based on an analogous concept to that of optical tweezers: a magnetic particle placed in an inhomogenous magnetic field experiences a net force in the direction of the gradient ($\vec{F} = -\vec{\nabla}U = 1/2\vec{\nabla}(\vec{m}(\vec{B}) \cdot \vec{B})$), where $\vec{m}(\vec{B})$ is the magnetization of the particle [23]. Consequently, using permanent magnets or electromagnets it is possible to exert forces to micrometer-sized magnetic particles and perform single-molecule manipulation experiments similar to those explained for AFM and optical tweezers setups.

Basic magnetic tweezers instruments are relatively straightforward to setup and consist of a pair of permanent rare earth magnets (>1.3 Tesla) placed above a microfluidics chamber that is mounted in an inverted microscope (Figure 5**a**) [4]. Functionalized super-paramagnetic beads ($D \sim 1$ μm) are generally used for the experiments, and molecules are tethered between the surface of the beads and the coverslip of the microfluidics chamber (Figure 5**a**). The force applied to the molecules is controlled by varying the distance from the magnets to the sample, and beads are imaged using a charge-coupled device (CCD) camera connected to a frame grabber.

Magnets are typically separated by a gap of 1 mm, and with the opposite poles facing each other. This sets a magnetic field that decays with a characteristic length of 1 mm, creating a trapping potential

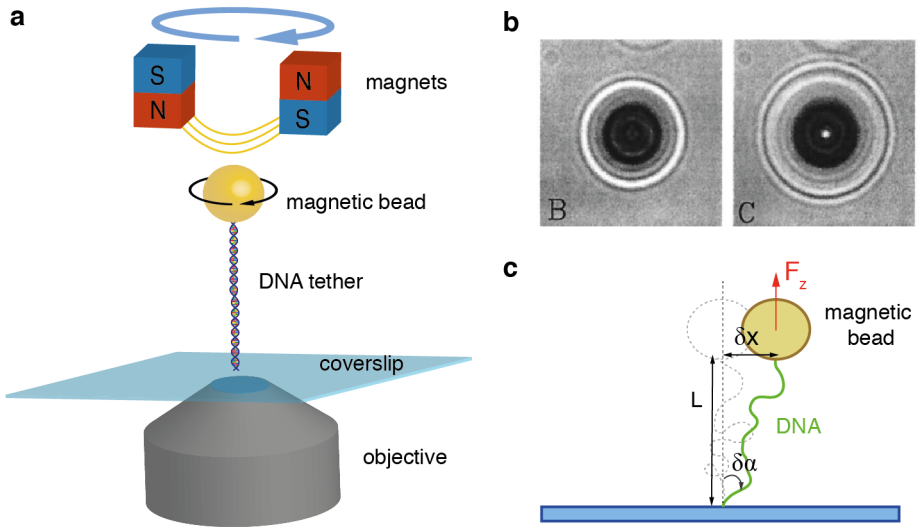

Figure 5: Magnetic tweezers. (**a**) Scheme of a magnetic tweezers setup. A DNA molecule is tethered between the coverslip surface and a magnetic bead. Force (torque) can be applied to the molecule by approaching (rotating) the magnets. Images of the tethered beads are obtained using an inverted microscope. (**b**) Interference pattern of a 4.5 μm bead at two different focal positions: 3 μm (left) and 5 μm (right) away from the surface. (**b**) Force can be calibrated from the lateral fluctuations (δx) of a magnetic bead tethered with a DNA molecule of length L.

with a stiffness on the order of 10^{-6} pN/nm [24]. This extremely low stiffness provides a surface of constant force to pull on the tethered molecules (i.e. a bead that moves its full size of 1 μm, only experiences a change in force of 10^{-3} pN). This makes magnetic tweezers very well suited to perform constant force experiments as they work as a passive force clamp, without need of a complex feedback system as for optical tweezers or AFMs. This has the advantage that constant-force experiments are not limited by the response rate of active feedbacks (typically below 1 kHz). On the other hand, the position of the beads cannot be controlled as for optical tweezers, and experiments are always performed using the force as the control parameter. The typical range of forces achieved with 1 μm sized beads is between 0.01-30 pN, although much higher forces (~ 200 pN) can be achieved by using large magnetic beads or electromagnetic tips [25, 26].

Another important feature of magnetic tweezers is the capability to exert torque. Super-paramagnetic beads contain thermally disordered magnetic domains that orient in the presence of an external magnetic field, providing a large magnetic moment in the direction of the field. However, beads are not fully isotropic, and have a preferential axis of magnetization that aligns with the magnetic field [4]. Consequently, beads can be rotated by turning the magnets, allowing to apply torque to the tethered molecules. Overall, the capability of magnetic tweezers to work as a natural force-clamp and exert torque, are the main features that have made them the instrument

of choice to study a large number of molecular motors and reactions that involve complex geometrical configurations such as transcription and replication [27, 28]. Similarly, the topological properties of DNA (e.g supercoiling), and the activity of enzymes that modify them (e.g. topoisomerases) have been largely investigated with this technique [29, 30, 31].

The position of the beads is usually tracked in real time from the CCD camera images. The beads create a diffraction pattern with many concentric rings due to the interference between the scattered and non-scattered light (Figure 5**b**) [32]. The lateral position is determined finding the centroid of the interference pattern, whereas the axial position is found by comparing its intensity profile with calibration images performed at a series of different focal distances. This calibration is done by moving the objective away from the beads by using a calibrated piezoelectric motor. As the interference pattern is a strong function of the distance between objective and beads, an axial resolution of 10 nm can be achieved (lateral resolution is ~2-4 nm) [4]. Force is directly determined from the distance between the magnets and the microfluidics chamber, and can be calibrated from the lateral fluctuations ($\langle x^2 \rangle$) of a bead tethered to the surface using a DNA molecule of known length (L). From the equipartition theorem follows that $1/2 k_B T = 1/2 k_x \langle x^2 \rangle$, where k_x is the lateral stiffness of the tethered bead. In this configuration, the stiffness on the transverse direction can be obtained by considering the bead-DNA system as an inverted pendulum of length L, where $k_x = F_z/L$ (Figure 5**c**). Bead position and force determination are therefore limited by the video-acquisition rate of the CCD camera (typically ~100Hz), but can be increased up to 1 kHz using most recent CCD technology [33]. Moreover, the use of high-speed cameras with a wide field of view allows parallel tracking of multiple beads (>100) [34, 35], yielding statistical data of a large number of molecules in a single experiment.

1.2.2 *Single-molecule fluorescence*

Another important set of tools to investigate biology at the single-molecule level are fluorescence-based techniques. Fluorescence techniques consist in the detection of light emitted by fluorophores that are attached to the molecules of interest. Fluorophores can be excited to a higher energy state by the absorption of light of a specific frequency, from which they relax back to the ground state through the emission of a lower frequency photon (Figure 6**a**). By combining illumination setups that excite very small volumes together with high-sensitivity detection schemes, it is possible to reduce background fluorescence and detect the light emitted by individual fluorophores. Very small volumes deep in a sample can be excited using confocal setups, and wide-field images of surfaces (~ 100 nm in depth)

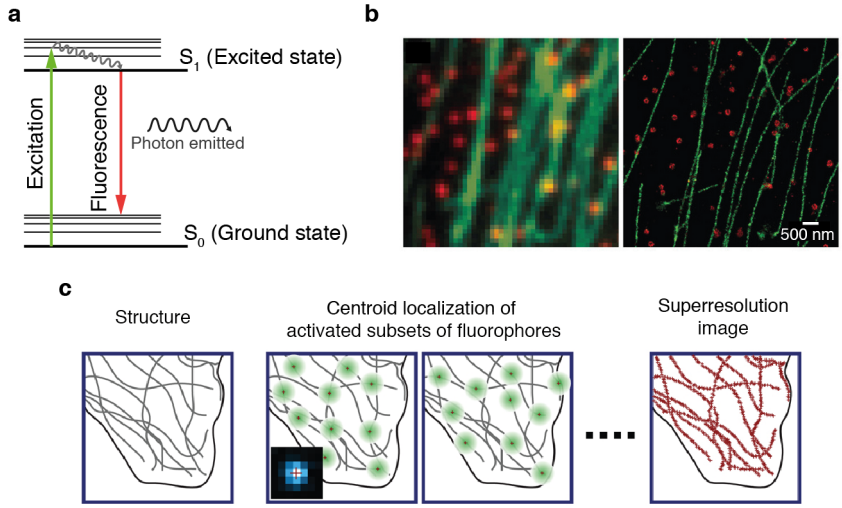

Figure 6: Single-molecule fluorescence and superresolution microscopy. (**a**) Jablonski diagram, showing the excitation of a fluorophore to its singlet excited state. The fluorophore relaxes first to its lowest vibrational level ($\sim 5-10$ ps) and then returns to the ground state through the emission of a lower energy photon ($\lambda_{emission} > \lambda_{absorption}$). (**b**) Comparison of a conventional fluorescence image and a superresolution image of a microtubules structure. (**c**) Superresolution images using photoactivable dyes. A structure smaller than the diffraction limit (left) is imaged by localizing the centroid of fluorophores that are sequentially activated in sparse subsets (middle). An image of a single dye (blue, inset) and its high-precision localization (red cross, inset) are shown. A superresolution composite image of the localized dyes can then be assembled (right). Adapted from [39].

are obtained using evanescent field illumination with total internal reflection fluorescence (TIRF) microscopy [36]. Overall, different approaches to single-molecule fluorescence allow to detect and follow the dynamics of a large number of biomolecular systems in real-time and with high temporal and spatial resolution, both in *in vitro* and *in vivo* conditions. Typical systems under study are conformational changes of proteins, protein-nucleic acid interactions, motor proteins or cellular membranes. A main advantage of fluorescence microscopy is its temporal resolution (on the order of μs for many setups), much higher than that achievable with force-spectroscopy techniques.

The maximum spatial resolution in optical systems is set by the diffraction limit ($\sim$ 300 nm). However, the centroid of the light spot created by a single fluorophore can be localized with a much higher precision that increases with the number of collected photons ($\sim 1/\sqrt{N}$) [37]. Superresolution microscopy techniques use centroid localization to overcome the diffraction limit, reaching resolutions up to 10-20 nm when a large number of photons ($N \sim 1000$) from a single emitting fluorophore are collected (Figure 6**b**) [38].

To accurately localize the centroid of the fluorophores the light spots created by different fluorophores in the sample must not overlap. However, in many experiments the molecules of interest are densely

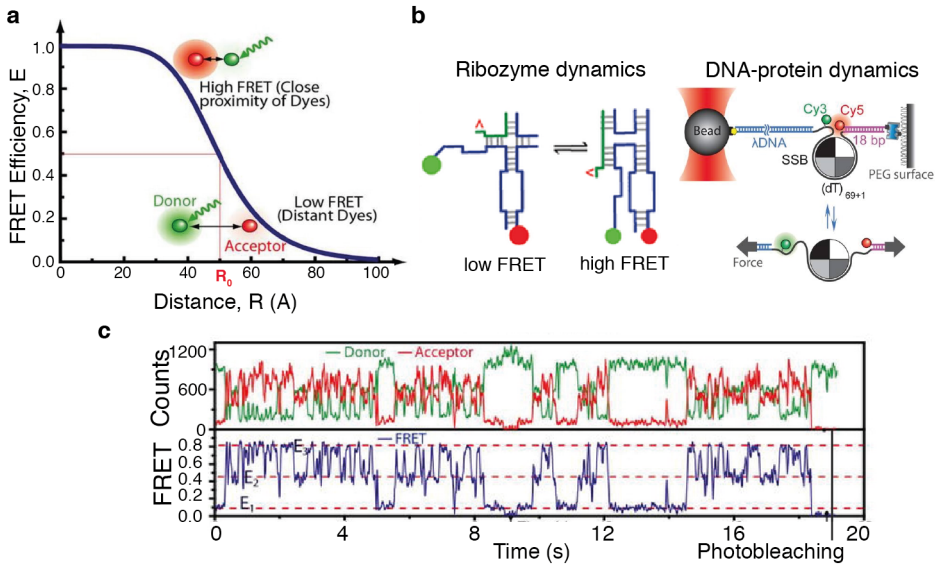

Figure 7: FRET experiments. (**a**) FRET efficiency as a function of inter-dye distance (R). A donor dye (green) excited with a laser beam either fluoresces (large interdye distance, $R > R_0$), or transfers the energy to the acceptor dye (red) if the dyes are close to each other ($R < R_0$). (**b**) FRET can be used to track fast conformational dynamics of molecules by placing the dyes at specific locations of the molecular structure. FRET setups can also be combined with force-spectroscopy techniques such as optical tweezers to follow conformational changes. (**b**) An example FRET trace for the ribozyme system shown in panel (b). The donor and acceptor fluorescence intensity (green, red) are anticorrelated due to the energy transfer associate to FRET. The FRET efficiency (blue) shows the existence of three conformational states characterized by different FRET efficiencies (E_1,E_2,E_3). The acceptor photobleaches after 18 s. Adapted from [44].

distributed (e.g proteins in a cellular membrane), and several fluorophores share a diffraction-limited volume. In these conditions the largest challenge becomes distinguishing individual fluorophores that overlap creating blurred images. In superresolution techniques individual detection in dense distribution of fluorophores is achieved either using (i) structured illumination techniques (e.g. STED) [40, 41] or (ii) photoactivable dyes that can be switched on and off with time (e.g. PALM, STORM) [42, 43]. Using photoactivable dyes, a composite high resolution image can be reconstructed from consecutive images in which different subsets of sparse fluorophores are sequentially imaged (Figure 6**c**) [42].

One of the most widely stablished single-molecule fluorescence techniques is fluorescence resonance energy transfer (FRET). In FRET measurements, the distance between two fluorophores (a donor and acceptor) can be measured in the scale of 1-10 nm (Figure 7**a**). By placing the two fluorophores at particular locations of a biomolecule, conformational changes can be followed at high temporal resolution (ms) (Figure 7**b**) [45]. In a FRET pair, the donor fluorophore is excited with a laser source of appropriate wavelength. If the acceptor

fluorophore is found at a very short distance (<10 nm) of the excited donor, part of this energy is transferred to the acceptor via a non radiative dipole-dipole interaction, that relaxes back to its ground state emitting a photon at another wavelength. This resonance energy transfer requires that the emission spectra of the donor and the absorption spectra of the acceptor partially overlap. The efficiency of transfer between donor and acceptor is then a strong function of the interdye distance in the nanometric scale: $E = 1/(1 + (R/R_0)^6)$ where R is the distance between donor and acceptor and R_0 the Förster radius at which the efficiency is 0.5 (Figure 7**a**). Consequently, FRET is an excellent technique to follow conformational dynamics and intramolecular reactions, as it provides an internal molecular frame of reference that strongly reduces sources of instrumental noise and drift (Figure 7**b,c**) [44]. Typically bright small organic dyes (< 1 nm), with good spectral separation such as the Cy3 and Cy5 pair are used for tagging. For most *in vitro* applications, the molecules of interest are typically immobilized in a surface and imaged using a TIRF setups in order to obtain high-throughput data [36].

Finally, recent approaches that combine a force-spectroscopy technique and single-molecule fluorescence are allowing to follow complex bimolecular reactions difficult to address with a single reaction coordinate [46]. For instance, visualization of intercalating dyes in the overstretching transition has unveiled the long-standing problem of the structure of DNA at high stretching forces [47, 48]. Similarly, the addition of a second reaction coordinate is likely to provide new information of processes that involve complex biomolecular structures such as DNA replication [27].

CHAPTER 2

OPTICAL TWEEZERS

Laser optical tweezers is a technique that uses a focused laser beam to exert an optical gradient force to particles that have a higher refractive index than the surrounding medium (e.g. plastic or glass microspheres in water). This optical force arises from the conservation of linear momentum, and in certain conditions, is strong enough to keep a particle trapped near the focus of the beam. It is then than an optical trap is created.

2.1 PRINCIPLES OF OPTICAL TRAPPING

Historically, the fact that electromagnetic radiation exerts a pressure on surfaces was already predicted by J.C. Maxwell [49], and proved experimentally at the beginning of the twentieth century in the Nichols-Hull radiometric experiment. However, the observation of a force generated by radiation pressure capable of moving micrometer-sized particles in the laboratory was not reported until 1970 by Ashkin, who took advantage of the high light intensities of laser technology to trap latex microspheres in water [50]. During the following years, the technique was developed until succeeding in creating stable optical traps able to confine dielectric microparticles using a single laser beam [51]. This achievement established the basis of what we currently know as optical trapping and paved the way for single molecule micromanipulation experiments, providing a novel tool to explore biology and physics at the frontier of Brownian forces.

So which are the forces that confine dielectric particles on a focused laser beam? In principle, Maxwell equations could be used to determine the electromagnetic forces that arise between particle and light in the experimental set-up. However, this is complex for most situations, and two approximations are generally used to obtain a qualitative insight:

- **Ray optics approach.** If the trapped particle is much larger than the wavelength of light (d>10λ), a ray-optics approach provides a qualitative interpretation of how an optical trap is created.
- **Rayleigh scattering conditions.** If the dielectric particle is much smaller than the wavelength of light, it can be considered as an electric dipole and the electromagnetic-field calculations are simplified.

Any of these approximations applies to the usual working conditions of most optical tweezers, as laser wavelengths are tuned in the near-infrared range to avoid damaging biological samples ($\lambda \sim 1$ μm) [52], and typical bead sizes are also in the order of 1-3 μm. Although these approximations are only strictly valid in the aforementioned limit conditions, they are useful to obtain a more intuitive idea of how this confinement is achieved without having to solve complex electromagnetic-force equations. Nevertheless, further explanation of the subject and force calculations can be found in [53] and [54].

2.1.1 *Ray-optics approach*

In the ray-optics approach we consider a dielectric particle much larger than the wavelength of the laser, and thus light beams can be tracked individually throughout the particle (Figure 8**a**, left). Light beams are refracted when they cross the boundary surfaces, and thus they suffer a change in direction. Since light carries a momentum $\vec{p} = \hbar\vec{k}$, this deflection also causes a change in this magnitude ($\vec{p}_{in} \neq \vec{p}_{out}$). Consequently, the particle experiences an equivalent force $\vec{F} = \frac{d\vec{p}}{dt}$ that pushes it towards the opposite direction due to linear momentum conservation

A vectorial decomposition of light momentum, helps analyze the reaction forces along its propagation direction (z), and its perpendicular plane (x-y):

- In the x-y plane, as the particle has a higher refractive index than the environment, light rays entering through the left side (Figure 8**a** left, ray 1) are deflected towards the right and thus push the particle to the left (f_1). On the other hand, light rays entering through the opposite side (Figure 8**a** left, ray 2) push the bead towards the opposite direction (f_2). However, these forces do not cancel each other, as laser beams generally have

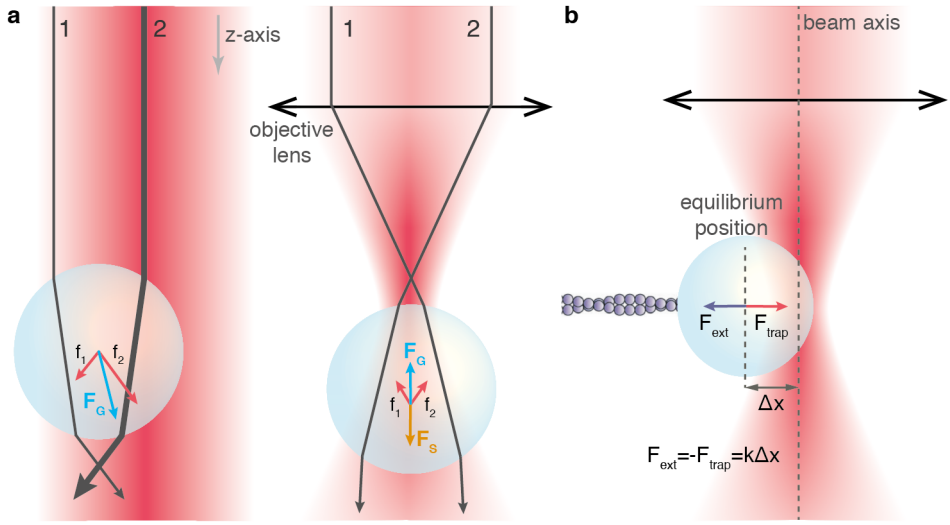

Figure 8: Optical trapping principles. (**a**) Ray optics diagram of a particle placed in a non-focused (left) and focused (right) Gaussian laser beam. Individual light rays are indicated with gray arrows (thickness proportional to intensity), and forces exerted by each ray are shown in red. For a focused beam, the addition of all light rays causes a net gradient force (F_G) that compensates the scattering force (F_S) allowing to stably trap the particle near the beam focus. (**b**) Optical tweezers as a force-transducer. When an external force (F_{ext}) is applied to a trapped bead, the bead is displaced from the beam axis. For small displacements (Δx) the trap behaves as an harmonic well. Therefore, if the trap stiffness k is know, the external force can be determined by measuring Δx.

a non-uniform intensity profile[1]. Hence, the more intense rays coming from the centre of the beam apply a higher force than rays coming from the edges, creating an effective restoring force that keeps the bead in the most intense region of the beam, just like a spring.

- In the propagation direction, the deflected light rays experience a decrease on the z-component of its momentum, and therefore the particle experiences a force that pushes it forward on the z-axis.

To stably trap the particle also in the z-axis, it is necessary to create an intensity gradient also in the propagation direction. This can be achieved by focusing the laser beam into a spot with an objective lens (Figure 8**a**, right). If a high NA objective is used, a steep light gradient can be created[2], and therefore a net restoring gradient force (F_G) towards the beam waist arises; effectively confining the particle in the 3-directions of space. Indeed the trapped particle reaches an equilibrium position slightly beyond the focal point of the laser beam. In

1 For optical trapping usually lasers with a Gaussian profile are used (TEM_{00} mode).

2 The diameter of the beam waist of a focused laser beam is $D \sim \frac{2\lambda}{n\pi NA}$, where λ is the light wavelength, n the difractive index of the medium, and NA the numerical aperture of the objective.

this position the gradient force (F_G) exactly compensates the scattering force (F_S) that points in the direction of the incident light. The scattering force is caused by those deflected light rays experiencing a decrease on the z-component of its momentum, as well as rays adsorbed or reflected by the trapped particle. Using a high NA lens is essential to confine the particle in the z-direction, as marginal rays (those coming from the edges) play a major role in creating the gradient force (Figure 8**a**, right). In fact, in single beam optical traps, it is common to use expanded laser beams that overfill a high NA objective in order to increase the intensity of marginal rays, and achieve a higher trapping force in the z-axis. Although this decomposition of the overall optical force in a scattering force and a gradient force is somehow arbitrary (as both forces arise from the same physical principles), it is usually considered as a convenient way to intuitively explain optical trapping.

2.1.2 *Electric dipole approximation*

A more quantitative insight can be obtained in the Rayleigh scattering limit. If the dielectric particle is much smaller than the wavelength of light (D«λ), it can be thought as an electric dipole placed in an inhomogeneous electromagnetic field. In this situation, Maxwell equations can be used to find a tractable expression that relates the trapping force to the light intensity gradient. The force in a charged particle placed in an electromagnetic field is described by the Lorentz equation [49]:

$$\vec{F} = q\left(\vec{E} + \frac{d\vec{r}}{dt} \times \vec{B}\right) \tag{3}$$

Thus the force experienced by an electric dipole in an inhomogeneous electric field can be obtained by adding the forces that the field applies to each of the two charges separated by an infinitesimal distance $d\vec{r}$:

$$\vec{F} = |q|\left((\nabla d\vec{r})\,\vec{E} + \frac{\partial d\vec{r}}{\partial t} \times \vec{B}\right) \tag{4}$$

This expression can be rewritten in function of the electric dipole moment ($\vec{p} = q\vec{r}$), and simplified considering that the dipole moment is also related to the electric field through the polarizability α ($\vec{p} = \alpha\vec{E}$):

$$\vec{F} = (\vec{p}\nabla)\,\vec{E} + \frac{d\vec{p}}{dt} \times \vec{B} \rightarrow \vec{F} = \alpha\left[\left(\vec{E}\nabla\right)\vec{E} + \frac{d\vec{E}}{dt} \times \vec{B}\right] \tag{5}$$

This expression can be simplified by using the vector identity $\left(\vec{E}\nabla\right)\vec{E} = 1/2\nabla\vec{E}^2 - \vec{E}\times\left(\nabla\times\vec{E}\right)$ and the Maxwell equation $\nabla\times\vec{E} = -\frac{d\vec{B}}{dt}$, finding that the force can also be expressed as:

$$\vec{F} = \alpha\left[\frac{1}{2}\nabla\vec{E}^2 + \frac{d}{dt}\left(\vec{E}\times\vec{B}\right)\right] \tag{6}$$

The magnitude $\vec{E}\times\vec{B}$ is directly proportional to the Poynting Vector $\vec{S}$, which is the energy flux of the electromagnetic field going through a surface. Under a steady applied force, the second term in Eq. 6 averages to zero:

$$\vec{F}(\vec{r}) = \langle\vec{F}(\vec{r},t)\rangle_T = \frac{\alpha}{2}\nabla\langle\vec{E}^2(\vec{r},t)\rangle_T = \frac{\alpha}{4}\nabla|\vec{E}^2(\vec{r},t)|^2 \tag{7}$$

Finally, as the intensity of a light beam is proportional to the square of the electric field ($I(\vec{r}) \propto (cn\epsilon_0/2)|\vec{E}(\vec{r})|^2$), it is straightforward to obtain that the force applied to the dipole is directly proportional to the light intensity gradient:

$$\vec{F}(\vec{r}) = \frac{\alpha}{2cn\epsilon_0}\nabla I(\vec{r}) \tag{8}$$

As a result, a dielectric particle placed in an inhomogeneous electric field experiences a Lorentz force in the direction of the intensity gradient. This force is proportional both to the light intensity gradient and the polarizability α of the electric dipole. Therefore, as the polarizability of a particle depends on its intrinsic properties (e.g. size, material, shape), different particles will experience different trapping forces even if the intensity gradient is the same. This is an important issue to take into account in some force-measuring systems, as calibration of the optical trap will only be straight-forward for plastic beads reproducible in size, and not for differently shaped particles or organelles.

2.2 FORCE MEASUREMENTS WITH AN OPTICAL TRAP

In the previous section we have seen that a laser can be used to stably trap micrometer-sized particles by means of intensity gradients, and we have used optical and electromagnetic theory to qualitatively explain where this trapping force arises from. However, the most exciting possibility of optical traps is not its unique ability to confine particles, but that they can also be used as a force-transducer to measure forces applied to the trapped object. Indeed, optical traps give access to forces in the piconewton (pN) range; in the limit of thermal fluctuations and in the scale of forces that drive most non-covalent biomolecular interactions.

2.2.1 *Force-measurements based on displacement determination and stiffness calibration*

A particle placed on a focused laser beam experiences a restoring force that holds it in an equilibrium position near the focal point. For small displacements, the optical trap behaves to a very good approximation as a harmonic potential well, and a Hookean spring approximation can be used ($F = k\Delta x$, where k is the characteristic stiffness of the trap and Δx the displacement from its center). Consequently, if the spring constant k is determined, the optical trap can be used to perform force-measurements just by precisely measuring the displacement Δx caused by the external force (Figure 8**b**). This force-measurement method, therefore, strongly relies in our ability to accurately determine these two magnitudes (k and Δx).

Position determination can be resolved for spherical beads with very high precision (sub-nanometer resolution) by means of several optical methods such as: algorithms for sub-pixel detection from CCD imaging (restricted by video acquisition rates) [55, 56]; direct imaging of the trapped object on a quadrant photodiode (QPD) [57]; back-focal plane interferometry [58]; or laser-based position detection (a single laser can be used both for trapping and displacement measurement). Discussion of these techniques can be found in most optical tweezers reviews [18, 59]. Regarding the calibration of the particle displacement, several methods have been developed, depending on the detection scheme and the ability to move the optical trap or the stage holding the microfluidic chamber. Generally, calibration is done by moving a particle fixed to the stage through a known distance, while recording the response on the displacement detector. Displacement calibration is only valid for other particles having the same shape, size, and orientation; and is also affected by changes in the axial direction (z-axis). Consequently, the displacement detector should be recalibrated for differently shaped particles; and as a result, the use of spherical particles of the same size is usually preferred [60].

Overall, this approach to determine the forces involved in an optical trap is the most commonly used at present, as it is a very well established methodology and relatively easy to set-up.

2.2.1.1 *Stiffness calibration*

The optical trap stiffness can be determined by several different methods. Two commonly used methods are measuring an externally controlled Stokes drag force ($\vec{F} = \gamma\vec{v} = 6\pi\eta R\vec{v}$ for a spherical bead), or performing Brownian motion force measurements (see section A.2.1). However, determining the trap stiffness by applying either Stokes law or thermal corner-frequency methods, requires good knowledge of the particle size and shape, and the local fluid viscosity. As for the displacement detector, this can be done to a good approximation

for spherical beads of known diameter, but not for particles of arbitrary size and shape such as cells and organelles [60]. Consequently, precise quantitative experiments with optical tweezers are generally performed in vitro and using spherical plastic beads for trapping. An alternative measurement of the trap stiffness, can be done by directly measuring the distance fluctuations of the trapped partice $\langle \delta x^2 \rangle$. From the equipartition theorem follows that each degree of freedom takes $1/2 k_B T$. This energy can be equated to that of the fluctuating trapped particle, finding that $k_{trap} = k_B T / \langle \delta x^2 \rangle$. This method has the advantage that is independent of the drag coefficient of the particle γ, but it requires recording the fluctuations with a frequency well above the corner frequency of the bead ($w_c = k_{trap}/\gamma$), which is on the order of a few kHz for typical setups. Indeed, it is usually advisable to measure the stiffness using the three different methods, as it might help to detect subtle problems with the performance of the optical trap [59]

2.2.2 *Force-measurement based on light-momentum changes*

The optical tweezers instrument used in this thesis, however, uses an alternative force-measurement method based on linear-momentum conservation [60, 61]. As we will see, this optical-trap force transducer overcomes some of the difficulties outlined before and has the advantage that force calibration becomes independent of most experimental factors.

2.2.2.1 *Underlying principle*

For a particle confined in an optical trap, and in the absence of external forces, there is not net transfer of momentum between the particle and the light beam, and therefore the beam is not deflected ((Figure 9**a**, top). However, when an external force is applied to the particle, this force is balanced by a change in momentum of the laser beam; or what it is the same, the laser beam is deflected (Figure 9**a**, bottom). As the net change in momentum in the light beam is equal to the applied external force ($\vec{F}_{ext} = \Delta\vec{p}$), just by comparing the intensity distribution of the light beam as it enters and leaves the trap, the external force can be determined independently of most experimental factors such as particle shape, size, refractive index and viscosity of the buffer, or variations in laser power [60]. This allows us to measure forces in irregularly shaped objects such as cells and organelles, and simplifies experiments with plastic microspheres as force is directly determined from first principles.

To properly estimate the force from the deflection of the laser beam, it is essential to fully recollect the deflected light beam once it exits the trap (so the incident and exiting light distributions can be compared). A way to do so is by placing a second objective lens on the

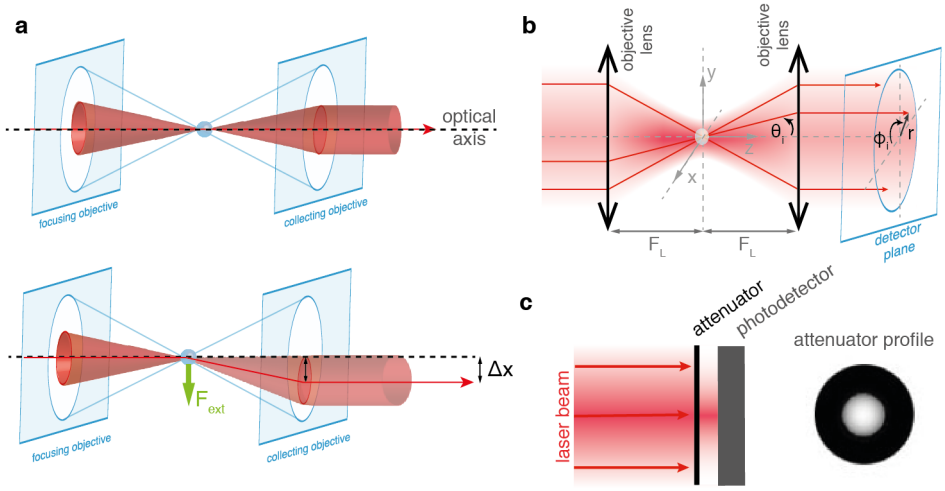

Figure 9: Force-measurement from light momentum. (**a**) Optical trap created with a low NA laser beam (red cone) and a high NA objective. In the absence of external forces (top) the laser beam is not deflected and remains on axis. External forces (bottom) can be determined by collecting the deflected laser beam with an objective lens, and measuring the distance shift Δx. In this set-up objective overfilling cannot be used as the deflected beam would exceed the maximum collection angle (blue cone). (**b**) Ray optics diagram for force determination. The laser beam is recollected with an objective lens placed a focal length away (F_L) from the trap. In this way, angles are converted to distances and projected into a PSD. (**c**) Photodetector used for axial force determination. An optical attenuator is used to measure the axial force from Eq. 12.

beam axis to collect the exiting laser beam. This method can not be applied straightforwardly in single beam optical tweezers, as it is very difficult to efficiently trap the particle and recollect the deflected exiting beam at the same time. In most single beam optical tweezers the incident objective lens must be overfilled with the laser beam in order to create a strong enough axial trapping force, making impossible to fully recollect the deflected exiting beam afterwards (fig.3). In fact, in this force-measuring method, a non-overfilling laser beam (low NA) and high NA objective lens must be used. By doing so, even when the exiting laser beam is deflected, it still reaches the collecting objective lens and the change in light momentum can be measured (Figure 9**a**, bottom). To reach high enough trapping forces in the axial direction in this non-overfilling configuration, two counter-propagating laser beams are used. In this way, the scattering force due to each laser is compensated by the opposite beam, and particles can be stably trapped in the axial direction. This set-up, that is the one used in this thesis, has the only drawback that requires a correct alignment of the two lasers, which is not always easily achieved.

2.2.2.2 *Force determination*

A ray of light can be understood as a beam of photons that carries a momentum flux $\frac{dP_i}{dt} = \frac{nW_i}{c}$, where W_i is the power carried by the ray, n the refractive index of the surrounding medium, and c the speed of light. If a light ray is deflected through angles θ_i and ϕ_i relatively to the optical axis (Figure 9**b**), then the force on the trapped particle due to the interaction with this ray is:

$$F_i = \frac{dP_i}{dt} = \left(\frac{nW_i}{c}\right)[\sin\theta_i \cos\phi_i \hat{x} + \sin\theta_i \sin\phi_i \hat{y} + \cos\theta_i \hat{z}] \quad (9)$$

As a result, the total force that the laser beam applies to the trapped particle can be found by adding up the force contribution F_i of all the light rays entering and leaving the trap, with the sign convention that un-deflected rays cancel each other (i.e. rays entering the trap have opposite sign than exiting rays) [61]. Moreover if the optical trap is placed at the focal plane of the collecting lens, light rays leaving the trap with an angle θ_i, exit the collecting lens parallel to the optical axis at a radial distance $r_i = n \sin\theta_i R_L$ (Abbe sine condition)[3]. Therefore, by placing the collecting lens a focal length away from the trap centre, the deflection angles θ_i become distances r_i that can be projected onto a photodetector [61].

An expression for the overall force in terms of the spatial intensity distribution $W(x, y)$ of the laser beam, can be found by combining Eq. 9 and the Abbe sine condition, and integrating over every light ray ($x = r\cos\phi$, $y = r\sin\phi$):

$$F_x = \left(\frac{1}{R_L c}\right)\left(\left[\sum W_i x_i\right]_{in} - \left[\sum W_i x_i\right]_{out}\right) \quad (10)$$

$$F_y = \left(\frac{1}{R_L c}\right)\left(\left[\sum W_i y_i\right]_{in} - \left[\sum W_i y_i\right]_{out}\right) \quad (11)$$

$$F_z = \left(\frac{n}{c}\right)\left(\left[\sum W_i \sqrt{1 - \left(\frac{r_i}{nR_L}\right)^2}\right]_{in} - \left[\sum W_i \sqrt{1 - \left(\frac{r_i}{nR_L}\right)^2}\right]_{out}\right) \quad (12)$$

The power deflections[4] can be directly measured by projecting the light beam that exits the collecting lens onto a position-sensitive photodetector (PSD) [61]. The current measured in the PSD[5] is proportional to the detector responsivity Ψ and to the power deflection in each di-

3 where R_L is the focal length of the collecting lens.
4 The distance-weighted sums of the light intensity (i.e. $\sum W_i x_i$ and $\sum W_i y_i$).
5 A PSD can be understood as a PIN junction photodiode that converts the incident light rays into x and y currents.

rection. The amplified X and Y current signals are therefore given by:

$$X = \Psi \frac{\sum W_i x_i}{R_D} \tag{13}$$

$$Y = \Psi \frac{\sum W_i y_i}{R_D} \tag{14}$$

where R_D is the half-width of the PSD detector [60].

Hence, the force in the transverse directions (F_x and F_y) can be directly determined from the change in the PSD signals (ΔX and ΔY) when a force is applied. As two counter-propagating beams are used in this force-measurement method, the total force in each direction is the addition of the force detected by each PSD (one for each laser beam):

$$F_x = \frac{\Delta X R_D}{c\Psi R_L} = \frac{(\Delta X_{laser1} + \Delta X_{laser2})\, R_D}{c\Psi R_L} \tag{15}$$

$$F_y = \frac{\Delta Y R_D}{c\Psi R_L} = \frac{(\Delta X_{laser1} + \Delta X_{laser2})\, R_D}{c\Psi R_L} \tag{16}$$

Here we assumed equal responsivities for both detectors. For the axial force (z-direction), the light intensity W_i is weighted by a power concentration function that falls off from the optical axis (equation 12). In other words, the force in the z-direction is related to the power concentration of the laser beam. Such a response can be obtained by placing an attenuator with the appropriate transmission profile[6] in front of a planar photodiode (Figure 9**c**) [61]. If the pattern radius r of the attenuator matches nR_L, then the detector response is:

$$Z = \Psi' \sum W_i \sqrt{1 - \left(\frac{r_i}{nR_L}\right)^2} \tag{17}$$

where Ψ' is the responsivity of the planar photodiode.

Hence the force in the axial direction can be again directly determined from the change in the photodetector signals Z (one for each laser beam). The only difference being that for the axial force these contributions must be subtracted due to the counter-propagating configuration:

$$F_z = \left(\frac{n}{c}\right) \frac{\Delta Z_{laser1} - \Delta Z_{laser2}}{\Psi'} \tag{18}$$

2.3 THE INSTRUMENT: MINITWEEZERS SET-UP

The optical tweezers used in this thesis consist of two counter-propagating 845 nm laser beams (200 mW laser diodes) focused into a spot in a

6 An attenuator having an intensity profile $\sqrt{1 - \left(\frac{r_i}{nR_L}\right)^2}$.

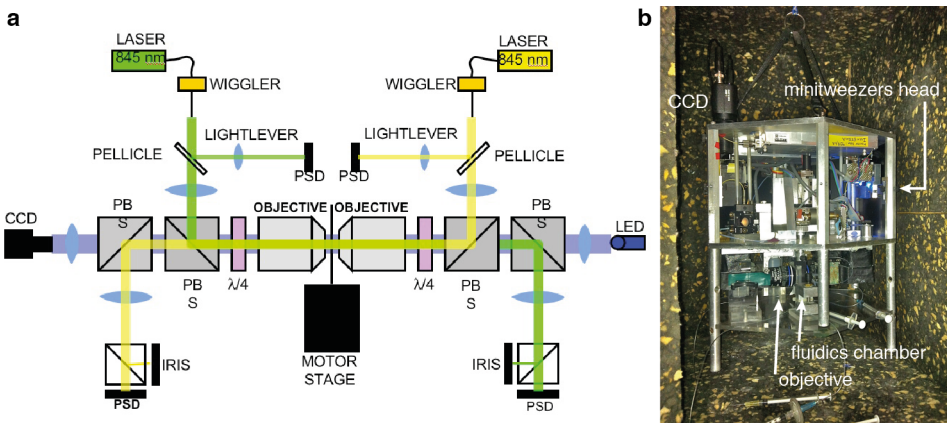

Figure 10: Minitweezers set-up. (**a**) Scheme of the instrument. The optical path of the counter-propagating lasers beams are shown in green and yellow. The position of the optical trap can be changed by means of piezoelectric actuators (wigglers), and the displacement measured by partially redirecting the beam onto light levers. Laser beams are focused onto a spot in a fluidics chamber where the optical trap is created, and then redirected to photodetectors (PSD, IRIS) to measure the force in the transverse and axial directions respectively. Images are obtained with a CCD camera. Reproduced from [62]. (**b**) Image of the minitweezers instrument without the protecting box of the head, where optics and detectors can be appreciated. Below, the objective lenses and the fluidics chamber are seen.

fluidics chamber by means of two high NA water-immersion objective lenses (Olympus UPlanSApo 60x/1.20). The laser beams can be deflected -and so the optical trap moved- by means of two piezoelectric actuators that are placed on the tip of the laser fibers. This allows to perform experiments with a higher accuracy than by moving the full microfluidics chamber, reducing vibrations and drift[7]. The piezoelectric deflectors have a range of $\sim$ 11 µm, sufficient for most experiments. However, the microfluidics chamber can also be moved by means of a motor stage for long distance displacements (e.g. trap beads) in a range of a few mm-cm.

A simplified scheme of the instrument and a brief description of the main components is presented in Figure 10**a**. The experimental set-up is composed of two counter-propagating laser beams (green and yellow) and uses a symmetric focusing and recollecting scheme for each laser beam. First, a pellicle beamsplitter is used to split the laser beam into a 96% component (trap beam) and a 4% component (reference beam). The reference beam is used as a light-lever, to precisely determine the position -and displacements- of the trap while performing experiments. The trapping beam is then directed to the microfluidics chamber using a prism box (PB)[8], and a focusing objective lens. A quarter-wave plate is placed before the focusing objective lens in order to obtain circularly polarized light. Circular polariza-

7 Drift effects can be further reduced by using a double trap configuration.
8 The PB contains a number of prisms and lenses for beam collimation and diverting

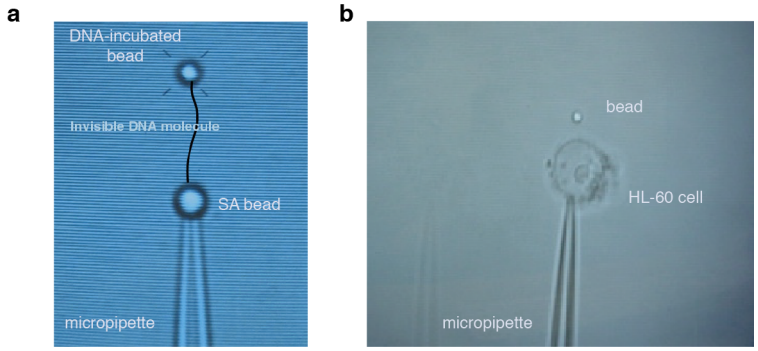

Figure 11: Experimental set-up. (**a**) DNA molecule tethered between two polystyrene beads. A streptavidin-coated bead is captured with a micropipette by air suction, and a DNA-incubated bead is subjected with the optical trap. (**b**) Human promyelocytic leukemia cell (HL-60 cell line) subjected with the micropipette. A polystyrene bead is trapped with the laser in order to perform membrane viscoelasticity measurements. The neutrophile precursor was previously trapped and displaced with the tweezers.

tion is important to obtain a spherically symmetric optical trap (on the x-y plane), and to differentially direct each laser beam to the corresponding PSDs for force measurements. The exiting laser beam is recollected by the opposite objective lens and becomes linearly polarized -in the transverse direction- by the opposite quarter-wave plate. Thanks to this differential polarization, the exiting laser beam can be redirected towards the PSDs by using a PB and a relay lens. For each laser beam two photodiodes are used for force measurement: one indicated as PSD, that is used to determine the force generated in the transverse direction (x-y plane), and another one indicated as IRIS, that is used to calculate the force in the axial direction (z-force). Finally, images of the fluidics chamber and the experimental set-up are obtained by means of a CCD camera, for which a blue led is used as the source of illumination using a Köhler scheme. The instrument (Figure 10**b**) has a very compact design (5 kg) and is based on the "minitweezers" design [60, 61].

2.4 EXPERIMENTAL CONFIGURATION

To manipulate individual molecules, micrometer-sized beads are coated with a chemical substance (e.g. streptavidin) that can specifically bind to its complementary molecule (biotin). Then, the molecule of interest (e.g. DNA, RNA) is biotinylated at one end, so it can bind through a strong non-covalent bond to the streptavidin-coated beads. To attach the opposite end of the molecule to another bead, a different specific bond is used to avoid attachments within a single bead. Typically, an antigen-antibody interaction is used due to its high specificity (e.g. digoxigen and anti-digoxigenin). Detailed synthesis protocols of the molecular constructs used in this thesis are found in Appendix D.

Experiments are then performed in a microfluidics chamber that contains a glass micropipette in which beads can be immobilized with air suction[9] (see Appendix B for details on the design and assembly of the microfluidic chamber). In a typical experiment (Figure 11**a**), a streptavidin-coated bead is captured with the optical trap and brought close to the micropipette, in which it can be immobilized by applying air suction. Then, a DNA-incubated bead is captured with the optical trap and approached to the pipette-immobilized bead. By bringing the two beads close to each other, a connection between the biotin in the DNA and the streptavidin in the bead can be made. This tether can then be pulled by moving the trap relatively to the micropipette; and force-extension curves measured. In another type of experiments, cells and organelles can also be directly trapped with the optical tweezers to carry out membrane viscoelastic measurements (Figure 11**b**).

9 Beads are differentially flowed into the microfluidics chamber through lateral channels. Typically, one channel is used to flow streptavidin-coated beads (D~2 μm), and the other channel is used to flow antidigoxigenin coated beads (D~3 μm) that have been pre-incubated with the DNA-labeled molecule.

CHAPTER 3

MECHANISMS OF SMALL LIGAND BINDING TO DNA

The mechanism by which small ligands bind to DNA has been a subject of intense research since the discovery of DNA structure nearly 60 years ago. Since the seminal work by Lerman, that coined the term "intercalation" to conceptualize the mechanism by which acridines bind to DNA [63], improved techniques (e.g. nucleic magnetic resonance [NMR], bioinformatics, X-Ray) have allowed to refine the structural features that govern the selectivity and affinity of DNA ligands. As the main focus of this thesis is the development of force-spectroscopy techniques to study the binding of small ligands to DNA, we will briefly review the current knowledge on their binding mechanisms. So far, two principal binding modes are generally used to classify the non-covalent association of small ligands with DNA: intercalation and minor groove binding. As we will see, these mechanisms can be further subdivided into different types (e.g. bis-intercalation, semi-intercalation), and indeed many ligands show combined mechanisms of binding that confers them higher selectivity for particular sequences and structures.

3.1 INTERCALATION

Intercalators are small planar compounds that bind nucleic acids by inserting between two adjacent base pairs (Figure 12). Consequently, intercalating drugs typically contain flat, aromatic ring systems that can fit between two contiguous base pairs, and whose insertion causes

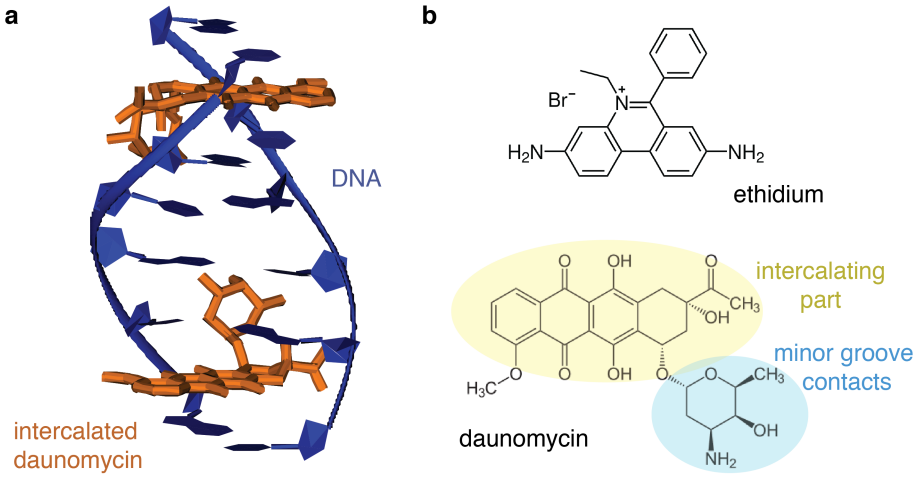

Figure 12: DNA intercalation. (a) Two daunomycin molecules (orange) intercalated in a DNA duplex with sequence CGATCG (blue). X-ray structure from PDB 1D11. **(b)** Chemical structure of two typical intercalators: ethidium and daunomycin. In daunomycin, the intercalating region is highligted in yellow, and a side group that makes contacts into the minor groove in blue. Both regions can be distinguished in the crystal structre shown in panel (a).

an unwinding and elongation of the DNA double helix. This physical distortion is the hallmark of DNA intercalation and is generally used to distinguish it from alternative binding modes [64, 65]. Although different intercalators vary in the degree to which they unwind the DNA double helix, they all elongate DNA about the same extent, increasing the contour length about 1 extra bp (i.e. ~ 3.4 Å) [66, 67].

Binding of an intercalator, requires breaking base-stacking interactions to create the intercalation cavity that accommodates the ligand between adjacent bases. It is generally thought that this initial free energy cost to establish the binding pocket, is compensated by the larger favorable contributions of the final intercalated state due to π-stacking between ligand and flanking base pairs, that is further stabilized by hydrophobic forces and additional intermolecular contacts (e.g. electrostatic, H-bonds) [68, 69, 70]. This distortion of the substrate upon ligand binding has been compared to a classical "induced-fit" model for enzyme-substrate binding [64]. Although intercalators show little sequence-selectivity, a generalized preference for pyrimidine-purine steps (5′-PyrPur-3′) has been observed [65, 71], and is attributed to the lower base-stacking interaction of these sequences that might facilitate its disruption and the insertion of the ligand [72]. Typically, the aromatic groups sit with their longest edge mostly parallel with respect to the base pairing axis to maximize stacking interactions. However it has been shown that intercalators that contain bulky groups at one side of their long axis, intercalate perpendicularly to the base

pairing axis [73, 74, 75]. For these ligands, flanking bases locally twist around the bound intercalator to maximize stacking interactions [75].

DNA intercalation was the first binding mode to be discovered for small ligands [63], and so far DNA intercalators constitute a large family of natural and synthetic products with applications that range from anticancer and antiviral agents (e.g. daunomycin) to DNA staining proves (e.g. ethidium bromide). Although these molecules share the insertion of planar compounds between base pairs, many of them also make additional contacts into the minor or major grooves, contributing to its stability and sequence selectivity [68, 65]. This is specially relevant for a second fundamental class of interaction, the so-called DNA bis-intercalation [76, 77].

3.1.1 *Bis-intercalation*

In a DNA bis-intercalator two intercalative moieties are connected by a linker chain. The compound is therefore able to "clamp" DNA by sandwiching two consecutive base pairs between its aromatic rings (Figure 13**a**); distorting and elongating DNA a length twice that of a mono-intercalator (i.e. $\sim 6.8 = 2 \times 3.4$ Å). The linker chain of a bis-intercalator usually sits on the minor groove -although binding via the major groove also exists [78]- and is an important determinant of the sequence-selectivity and affinity of the ligand [79, 77, 80].

Bis-intercalators were first discovered as naturally occurring peptides that had potent antiviral and anticancer activities, being echinomycin the paradigmatic and more widely studied one [81, 77, 82, 76, 83]. Natural bis-intercalators usually have uncharged bulky peptidic chains that establish a number of H-bonds with the DNA bases through the minor groove (Figure 13**b**). These contacts, and not the intercalating rings are considered to be the main determinant of their sequence-selectivity [64]. For instance, echinomycin and triostin A specifically clamp CpG steps, whereas sandramycin preferentially binds ApT steps [81, 77, 65].

The concept of bis-intercalation, was later used to create DNA staining proves with binding affinities much larger than typical monointercalating dyes. To create these optimized *de novo* fluorophores, two intercalating cromophores (e.g. ethidium, YO-PRO1) were bridged through an organic cationic linker (Figure 13**b**) to produce compounds that form stable, highly fluorescent complexes with dsDNA such as YOYO-1 or TOTO-1 [84, 85]. In this case, the linker chain increases the binding affinity due to the electrostatic attraction between cationic amino groups on the chain and the negatively charged DNA phosphate backbone, and through van de Waals contacts [79]. This electrostatic interaction with the phosphate backbone is also likely related to their little sequence-selectivity and comparatively faster binding kinetics. These dyes bind with a high affinity to DNA and show fluo-

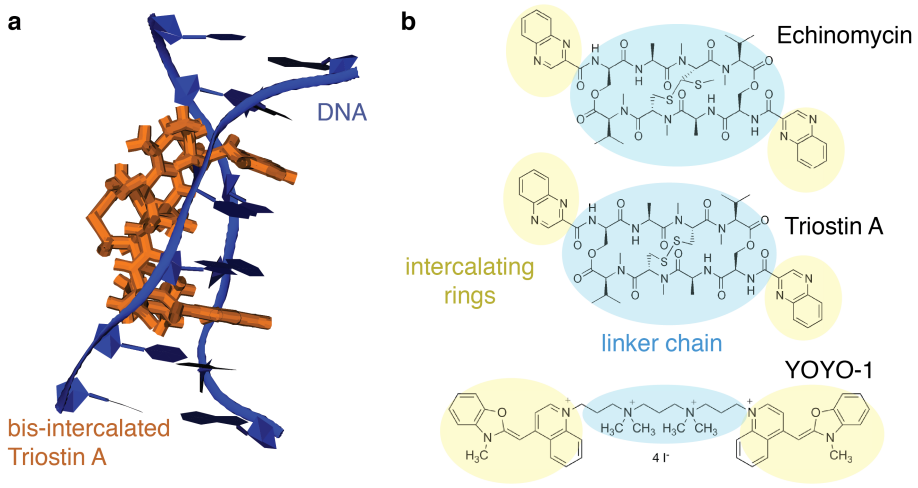

Figure 13: DNA bis-intercalation. (a) A Triostin A molecule (orange) bis-intercalated in a DNA duplex with sequence GACGTC (blue). X-ray structure from PDB 185D. **(b)** Chemical structure of 3 typical bis-intercalators: echinomycin, Triostin A and YOYO-1. All of them bind DNA through the minor groove. Note the difference in linker chain size between the two natural peptidic bis-intercalators and the organic dye (YOYO-1).

rescence enhancements >1000-fold upon binding [85], and have thus become the fluorophores of choice for a large number of biotechnological applications [86, 87].

3.1.2 *Threading intercalation and polyintercalators*

Many intercalators contain substituents of varying chemical nature at one side of the intercalating aromatic rings. Usually, these bulky groups cannot sterically fit within base pairs and sit at the minor or major groove of the DNA. In a threading intercalator, bulky side groups are found at opposite sides of the intercalating ring system, and upon binding one lies in the minor groove while the other interacts with the major groove [88, 89]. Consequently, in order to reach the final intercalated state, one of the bulky groups must pass through the DNA base pairs. This process is very different from the insertion of planar compounds as it requires a transient base pair destabilization (i.e. local DNA melting or breathing); making both the association and dissociation kinetics of these ligands extremely slow [90, 91]. Consequently, although threading intercalators show similar affinities than classical intercalators their kinetic rates are markedly reduced, providing tools able to interfere with biological processes for prolonged periods of time (e.g. blocking transcription factors). Among threading intercalators, ruthenium complexes have been particularly studied due to their chemical versatility [75, 91], and tunable photo-

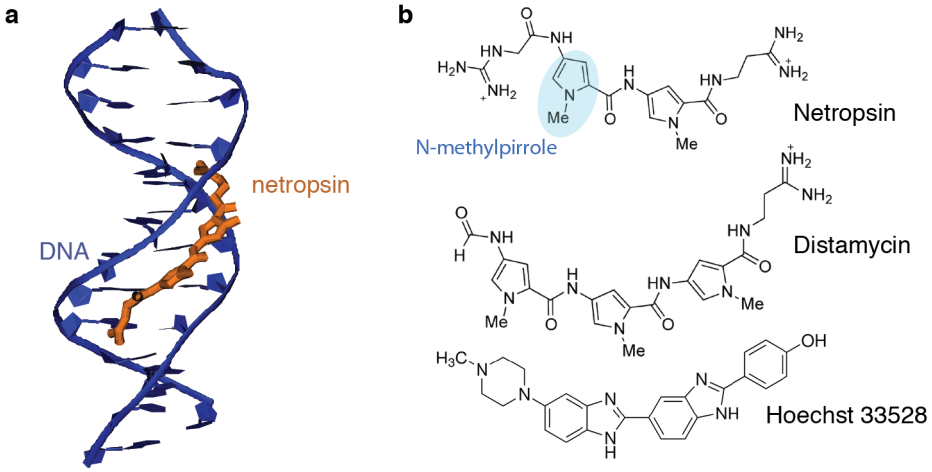

Figure 14: Minor groove binding. (a) A Netropsin molecule (orange) bound in the minor groove of a DNA duplex with sequence CGCGATATCGCG (blue). X-ray structure from PDB 1VTJ. **(b)** Chemical structure of 3 typical minor groove binders: netropsin, distamycin and Hoechst 33528. All of them share a crescent-shaped structure that fits well into the DNA minor groove.

physical properties, that can provide useful photoactivated biological applications[1] [92].

Finally, as threading intercalators protude through both DNA grooves, they allow to design multi-intercalating ligands in which the linker chain alternatively goes through the minor and major grooves in a snake-like fashion [93, 94]. For this sake, polyamide linker chains that bind the minor and major groove with sequence-selectivity have been used [95, 96], creating complexes that show exceptionally slow kinetic rates of up to 16 days and sequence-selectivity up to 14-bp [97].

3.2 MINOR GROOVE BINDING

Minor groove binders generally have a curved structure, and bind DNA by insertion into the minor groove without inducing a deformation or perturbation of the DNA structure (Figure 14**a**) [64, 96]. Their binding has therefore been compared to a "lock-and-key" mechanism, in which the ligand acts as a shape-selective binder that matches the helical pitch of the minor groove (Figure 14**b**) [64]. They were discovered well after intercalators, and have found applicability as antiviral compounds (e.g. distamycin, netropsin), and for *in vivo* fluorescence (e.g. Hoechst 33342).

These ligands are usually cationic, and their binding is stabilized by the formation of non covalent contacts (H-bonds, van der Waals, hydrophobic forces) with the nucleobases and the deoxyribose rings [98, 99]. In fact, most naturally occurring minor groove binders have a se-

1 For instance the photoactivation of an inert complex into a cytotoxic drug for cancer photodynamic therapy.

lectivity for the narrower and deeper minor groove found in stretches of AT base pairs[2] [99]. For instance, the aromatic groups found in netropsin and distamycin (N-methylpirrole, see Figure 14**b**), stablish H-bonds to the base edges of AT base pairs (particularly, with N3 of adenine and O2 of thymine) [69, 99]. In the quest to develop sequence-specific ligands, analogues of N-methylpirrole specifically tuned to discriminate the 4 possible base pair configurations (AT, TA, CG, GC) through the minor grove have been synthesized (Figure 15**a**) [96]. A polyamide strand can been used as a scaffold to link several of these heterocycles together in a modular fashion, and so create synthetic polymeric ligands able to specifically target predetermined DNA sequences through the minor groove (Figure 15**a**) [96]. As previously explained, this polymeric minor groove binders have also been exceptionally useful to link monointercalators or threading intercalators together and create high affinity, sequence-selective DNA binders [97].

3.3 OTHER BINDING MECHANISMS

3.3.1 *Major groove binding*

The major groove is richer in donor and acceptor sites for H-bonding, and most proteins make contacts to read this information and achieve very high binding selectivities [100]. However, small ligands that bind through the major groove are very uncommon. This lower abundance seems to be due to the wider and shallower shape of the major groove, which might not create an appropriate pocket for small ligand binding [65]. In fact, most small ligands that bind to the major groove also have an intercalating part that further stabilizes the complex. For instance, in the bis-intercalator ditercalinium, contacts on the major groove provide selective binding for GpC steps, but most of the binding free energy comes from the insertion of the intercalating rings [101, 78].

3.3.2 *Electrostatic interactions*

Electrostatic interactions with the sugar-phosphate backbone stabilize the binding of many small ligands that bind DNA through the previously described mechanisms. Many of these drugs contain cationic groups that neutralize the phosphate charges inducing the release of condensed counter ions (e.g. Na^+ and Mg^{2+}), which provides and entropic contribution to the free energy of binding [102, 103].

2 The width of the minor groove can vary from 3-4 Å in AT stretches to 8 Å in CG stretches. Its depth (8 Å in average) is shallower in CG regions due to the protruding C2-amine found in guanine that has no counterpart in the minor groove side of adenine or thymine.

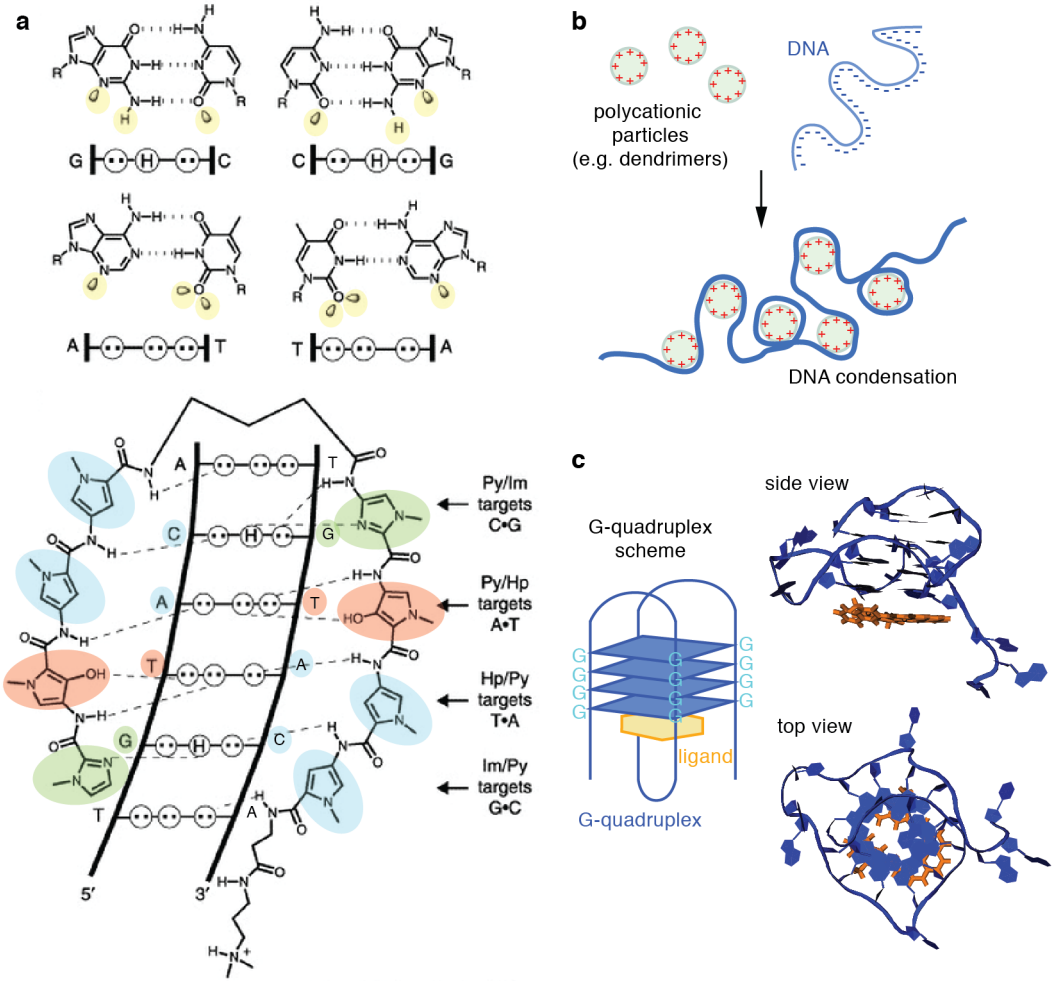

Figure 15: Sequence-selectivity and alternative binding modes. (a) Top: The minor groove shows a distribution of H-bond acceptor and donors (yellow) that allow to distinguish the 4 possible base pair configurations. A scheme of H-bond donors (H) and acceptors (lone pairs, two dots) are indicated below each base pair. Bottom: Derivatives of methylpirrole linked with a polyamide backbone can selectively bind DNA sequences by H-bonding with each base. N-methyl-pyrrolle recognizes C and A (blue), N-methyl-imidazole binds G (green) and N-methyl-3-hydroxypyrrolle binds T (red). **(b)** Polycationic particles (green) non-specifically bind DNA (blue) through the negative phosphate backbone inducing its condensation and compaction. **(c)** Targeting of G-quadruplexes. Scheme (left) and X-ray structure (right) of a G-quadruplex bound to a ligand (Phen-DC3) that increases the stability of the complex due to stacking interactions. X-ray structure from PDB 2MGN.

Although such interactions contribute to the binding free energy of many small ligands -and of proteins-, molecules that bind DNA exclusively by interacting with the DNA backbone are rare [104, 69]. This is due to the fact that these contributions are relatively small for mono cationic molecules and that they are strongly salt-dependent (i.e. they decrease with increasing salt concentration). An exception are polycationic polymers used for DNA condensation and gene transfer [105, 106], and ligands that can self-associate to form higher order aggregated structures that stack on the DNA phosphate backbone (Figure 15**b**) [69, 107]. Recently a drug designed to create a phosphate clamp through H-bond interactions has been developed and showed to associate with DNA exclusively with electrostatic interactions with the phosphate backbone [108].

3.3.3 *Targeting specific structures*

Finally, the implication of particular DNA structures (e.g. mismatches, G-quadruplexes, Z-DNA) in the mechanisms of gene regulation and carcinogenesis has also driven interest in developing ligands that particularly target these motifs [109, 110]. For instance, regions that can interconvert between G-quadruplex[3] and ssDNA are over-represented in telomeric regions and promoters of oncogenes [111]. Consequently, small ligands that stabilize G-quadruplexes and interfere with telemorerase activity or expression of oncogenes, hold a high potential as selective anticancer agents (Figure 15**c**) [112, 113]. Similarly understanding how the accessibility of DNA sequences in nucleosomes affects the biological activity of DNA binding drugs, is becoming a promising field to better understand their mechanism of action and develop more effective drugs [114].

3 G-quadruplexes are strucures that form in guanine-rich regions due to the stacking of consecutive guanines that fold into a four-stranded structure.

CHAPTER 4

MECHANISMS OF AGGREGATION

Several human disorders arise from the inability of certain peptides or proteins to remain in their native functional state. A large number of these protein misfolding diseases, end up with the conversion of non-native soluble species into aggregates of peptides or proteins that become insoluble when they exceed a certain size. A main focus of this thesis has been to develop tools to investigate the DNA-binding properties of anticancer peptides with limited solubility that aggregate even at low concentrations (μm). Indeed in part II, the aggregation and complexation of an anticancer agent (Kahalalide F) with nucleic acids is fully characterized and preliminary results with amyloidogenic peptides are presented. Consequently an introduction on the mechanisms that drive molecular aggregation and self-assembly, and their biomedical implications follows.

4.1 AMORPHOUS AND STRUCTURED AGGREGATES

Under appropriate external conditions most proteins self-assemble into aggregated particles, which can be conveniently classified between ordered and disordered aggregates [115, 116]. Amyloid fibrils are a clear example of ordered aggregates in which highly structured filaments are formed due to the self-assembly of peptides or proteins into protofilaments (of about 2-5 nm in diameter) that either twist together to form larger rope-like structures or associate laterally to form ribbon-like structures (Figures 16 and 18) [117]. On the other hand inclusion bodies and aggregates of α—crystallin in patients with

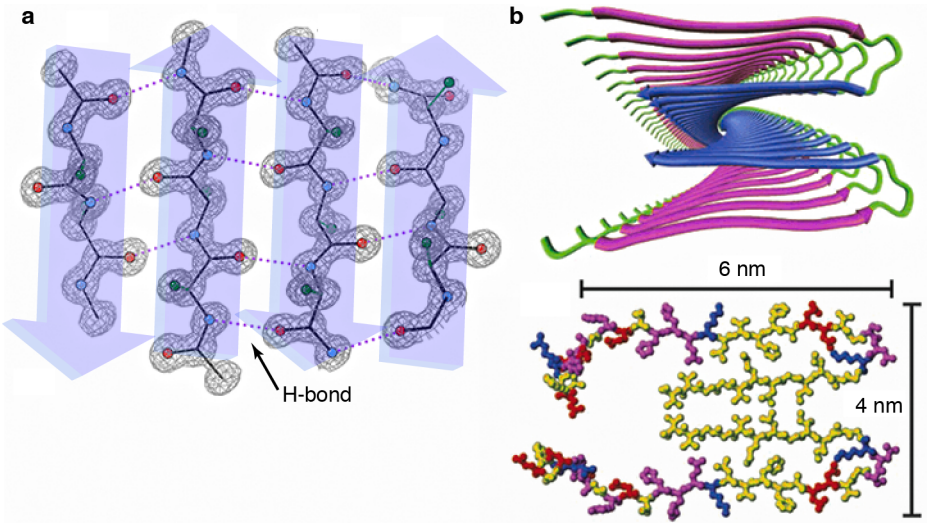

Figure 16: Amyloid structure. (a) β-sheet is a hallmark of amyloid. Antiparallel β-sheet showing how H-bonds between NH groups (blue) and CO groups (red) in the backbone of adjacent strands stabilize β-sheets. Side groups stick out of the plane (green). **(b)** Structural model of Aβ-40 assembly from ssNMR. (Top) Each peptide contains two β-strands (magenta, blue) that assemble into two parallel β-sheets. A protofilament (a four-layered structure) is seen down the long axis. (Bottom) The protofilament is stabilized by contacts between β-sheet layers due to side chain-side chain interactions, which are largely of hydrophobic origin. Color coding shows residues with hydrophobic (yellow), polar (magenta), positive (blue) and negative (red) side chains. Adapted from [118].

cataracts are examples of disordered aggregates, where proteins are thought to randomly self-assemble through hydrophobic interactions of partially unfolded or denatured proteins [115].

4.1.1 *Structured aggregation: Amyloid formation*

4.1.1.1 *Amyloid structure and driving forces of self-assembly*

The structure of amyloid fibrils has been widely studied through X-ray crystallography and solid-state NMR. It is known that amyloid fibrils are composed of a number of individual protofilaments (usually 2-6) that self-assemble into the final fibrils by twisting together or binding laterally [117]. Experiments using the previous techniques have shown that proteins forming each protofilament are arranged in β-strands oriented perpendicular to the long axis of the fibril, creating β-sheets that run in the direction of the fibril (Figure 16).

Although conformational polymorphism is generally found in amyloid fibrils, the high degree of uniformity observed in some of these structures is reminiscent to that of crystalline materials. In fact, and as we will later see, several aspects of amyloid formation are analogous to a crystallization process, whereas the formation of amorphous aggregates can be compared to a glass transition [116, 119,

120].Typically, the presence of amyloidogenic tissue is detected from the following hallmarks: Congo red birefringence and thioflavin T fluorescence as amyloid fibrils specifically bind these dies; extensive β-sheet structure observed in X-Ray crystallography, circular dichroism (CD) or infrared spectroscopy; and fibrous morphology in transmission electron microscopy (TEM) or AFM imaging. Despite the implication of very specific proteins in amyloid diseases, increasing evidence supports the idea that amyloid and its characteristic β-fold is a general structural form of polypeptide chains, and that virtually all proteins can self-assemble into amyloid fibrils [121, 122]. This comes from the discovery that many proteins that can form amyloid fibrils do not have common amino-acid sequences (i.e. sequence homologies) nor relevant structural similarities in their native fold [123]. These recent discoveries, have taken to the current paradigm that formation of amyloid takes place when the native conformation of a peptide or protein is destabilized under conditions in which other non-covalent interactions such as hydrogen bonds (essential for β-sheet formation) are favored. Generally, this happens through the formation of partially unfolded intermediates, in which certain regions of the polypeptide chain are able to nucleate the formation of β-sheets. In such conditions, the susceptibility of the polypeptide chain to end up as an amyloid fibril is mainly influenced by potential side chain-side chain interactions that can either stabilize or destabilize the β-strands. Among them, three types of interaction outstand as key factors affecting the rate at which partially unfolded proteins aggregate into amyloid fibrils [117]:

- **Hydrophobicity of the side chains.** The relevance of hydrophobic interactions in amyloid formation can be easily appreciated in the assembly of Aβ-40 protofilaments (Figure 16**b**). Indeed, the rate and propensity to amyloid formation of several proteins can be reduced (or increased) by performing point mutations that decrease (or increase) the hydrophobicity of the protein regions that nucleate aggregation [124, 125]. This also explains why long runs of hydrophobic residues are rarely found in highly expressed soluble proteins (where aggregation of partially unfolded intermediates is undesirable) [126], and the fact that aggregation rates of several soluble proteins have been found to be anticorrelated to their *in vivo* expression levels (i.e. the higher the expression level required for normal functioning, the lower their aggregation propensity) [127].

- **Polypeptide net charge.** It has been shown that there is an inverse correlation between the aggregation rates of proteins variants and their overall net charge (i.e. the higher the overall charge of the mutated protein, the lower its aggregation rate) [128]. This is due to the fact that a high net charge reduces protein aggregation due to an increased electrostatic repulsion

between monomers. Moreover this resonates with the finding that most amyloid deposits are found *in vivo* associated to negatively charged macromolecules (e.g. polysaccharides, nucleic acids), which enhance aggregation due to electrostatic attraction [129, 130, 131, 132].

- **Secondary structure propensity of the chain.** In addition to the previous, another key factor facilitating amyloid formation is a high propensity to β-sheet structure and a low propensity to α-helical structure; for instance due to alternate patches of hydrophobic and hydrophilic residues in the protein sequence.

4.1.1.2 *Kinetics of formation follow a nucleation and growth process*

These findings have permitted to rationalize and predict the rates of amyloid formation of different proteins and their mutants on the basis of this set of basic physicochemical properties (hydrophobicity, charge and secondary structure propensity) [133]. More broadly, other environmental factors that affect the rate of fibril formation are those able to promote partial unfolding of the proteins (e.g. high temperature, low pH) or interfere with the previous driving forces (e.g. high ionic strength). Overall this direct relation between simple physicochemical properties of partially unfolded polypeptides and aggregation rates indicates that amyloid formation is more similar to a polymerization or crystallization reaction than to the folding of globular proteins (where the rates of folding to the native state are highly dependent on the full individual sequence and domains of the protein) [117]. In fact, it is now widely accepted that amyloid fibril formation proceeds through a nucleation and growth process similarly to that observed in most polymerization reactions (e.g actin, tubulin) [134].

Nucleation and growth processes have been well studied in polymer physics [135]. They are characterized by an initial "lag phase" in which nucleation centers are created (e.g. protein oligomers), that is followed by a rapid exponential growth phase in which further monomers or oligomers are added to each preformed nucleus (Figure 17**a**). This reaction proceeds until there are no more free monomers available in the solution and saturation is reached. Hence, this generally leads to a characteristic sigmoidal behavior on the kinetics of these reactions (Figure 17**b**).

From a thermodynamic point of view, the formation of a nucleus is achieved when the gain in energy of adding further monomers to the aggregate is much larger than the loss of translational entropy of removing the monomers from the solution. Hence, in many of such processes the formation of the initial nuclei requires the cooperative aggregation of several monomers to reach this critical size, leading to a highly concentration-dependent all-or-none mechanism (i.e. a mini-

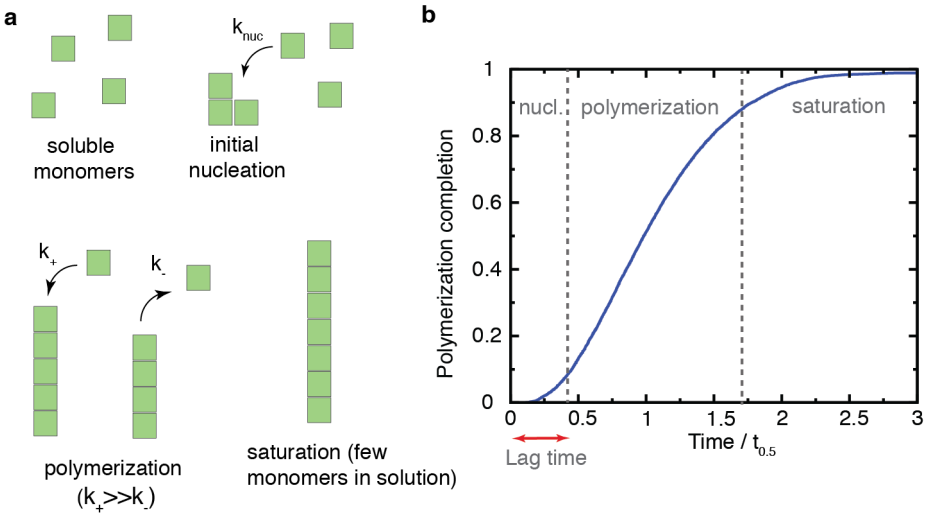

Figure 17: Nucleation and growth processes. (a) Scheme showing the steps involved in a nucleation and growth process: the initial formation of a nucleation center that requires the cooperative assembly of several monomers (driven by a rate of nucleation k_{nuc}); and the following extension of a formed nucleus (driven by a large rate of polymerization k_+) that proceeds until no free monomers are found in solution and saturation is reached. **(b)** Simulation of a nucleation and growth process showing the characteristic sigmoidal profile of such reactions.

mum critical concentration of monomer or protein is required to trigger fibrillization). This causes a large variation on the timescale of the initial lag phase required for nuclei formation depending on the specific conditions of the experiment (e.g. for the peptide Aβ-40 the lag time can vary from seconds to weeks depending on parameters such as concentration, ionic strength and pH). This all-or-none response is characteristic of most aggregation processes, and plays a major role in the mechanism of action of amorphous aggregating peptides such as the anticancer drug Kahalalide F studied in this thesis. As a general remark, a main challenge in single-molecule experiments with aggregating systems (such as those performed in this thesis with peptides Kahalalide F and preliminarily with Aβ-40) is therefore to set the conditions in which activity (i.e. complexation and aggregation) can be observed in the course of an experiment, but without having reached saturation (i.e. setting conditions with a lag time and kinetics compatible with single-molecule manipulation). Finally, it is noteworthy that quantitative models of nucleation and growth that account for secondary nucleation pathways (e.g. fragmentation of filaments, formation of branches within filaments) have recently proven very useful to describe the kinetics of amyloid formation *in vitro* [136], showing how quantitative models from physics combined with state-of-the-art experiments can help to elucidate unifying principles underlying complex biological and biomedical problems such as protein aggregation.

4.1.1.3 *Early oligomeric species and relation to disease*

From a biomedical perspective, understanding the kinetics of formation of amyloid fibrils is of particular interest as there is increasing evidence that are precisely these initial protein oligomers, those likely to be the pathogenic agents of most protein deposition diseases [137, 138, 139]. In fact, it has been proposed that the formation of the mature fibrils themselves might have a protective function by creating a depletion of the prefibrillar toxic species [140]. Consequently, during the last years great effort has been placed in elucidating the structure and kinetics of the intermediate species that precede the formation of the final amyloid fibers.

The pathways by which amyloidogenic proteins might fold-up into their functional form or self-assemble into structured or disordered aggregates shows a remarkable complexity (Figure 18). The general view is that when a protein is translated by a ribosome, it usually exits the complex as a largely unfolded polypeptide chain that collapses into a *molten globule* structure in which exposure of hydrophobic residues is minimized. Subsequently, the different domains of the protein fold into their secondary and tertiary structures, going through a series of intermediate partially unfolded states, until the final native conformation is reached. In some cases such proteins do self-assemble into functional oligomers (quaternary structure) or into large polymeric chains to accomplish their function (e.g. microtubules, actin filaments). The structures created in this canonical view appear from left to right in Figure 18. As explained in the previous subsection, on their way from an unfolded polypeptide chain to the final functional state, proteins are especially susceptible to self-assemble either as disordered aggregates, or as partially unfolded aggregates. Such oligomers might contain nuclei that seed amyloid (e.g. β-sheet structure) and be on-pathway towards the formation of a fiber (bottom of Figure 18); or form off-pathway unstructured aggregates (top of Figure 18) that might eventually dissociate again. Cells have evolved mechanisms to remove such species (e.g. by assisting folding with chaperones, or through degradation and recycling of misfolded proteins by the proteasome), but in some circumstances aggregation overcomes such control mechanisms leading to disease.

Consequently there is a large interest to identify and characterize these transient oligomeric species likely responsible of several neurodegenerative disorders. A well studied case is that of Aβ-peptides (particularly Aβ-40 and Aβ-42), that are a proteolysis products of the amyloid precursor protein (APP), and are the major fibrillar component of amyloid plaque in Alzheimer's disease [140, 141]. Several studies have suggested that the formation of Aβ-42 oligomers is the pathogenic event in Alzheimer's disease [142, 143, 144, 145]. AFM imaging and TEM studies have shown that Aβ-peptides can assemble to form oligomers with various shapes: annular pore-like structure, as

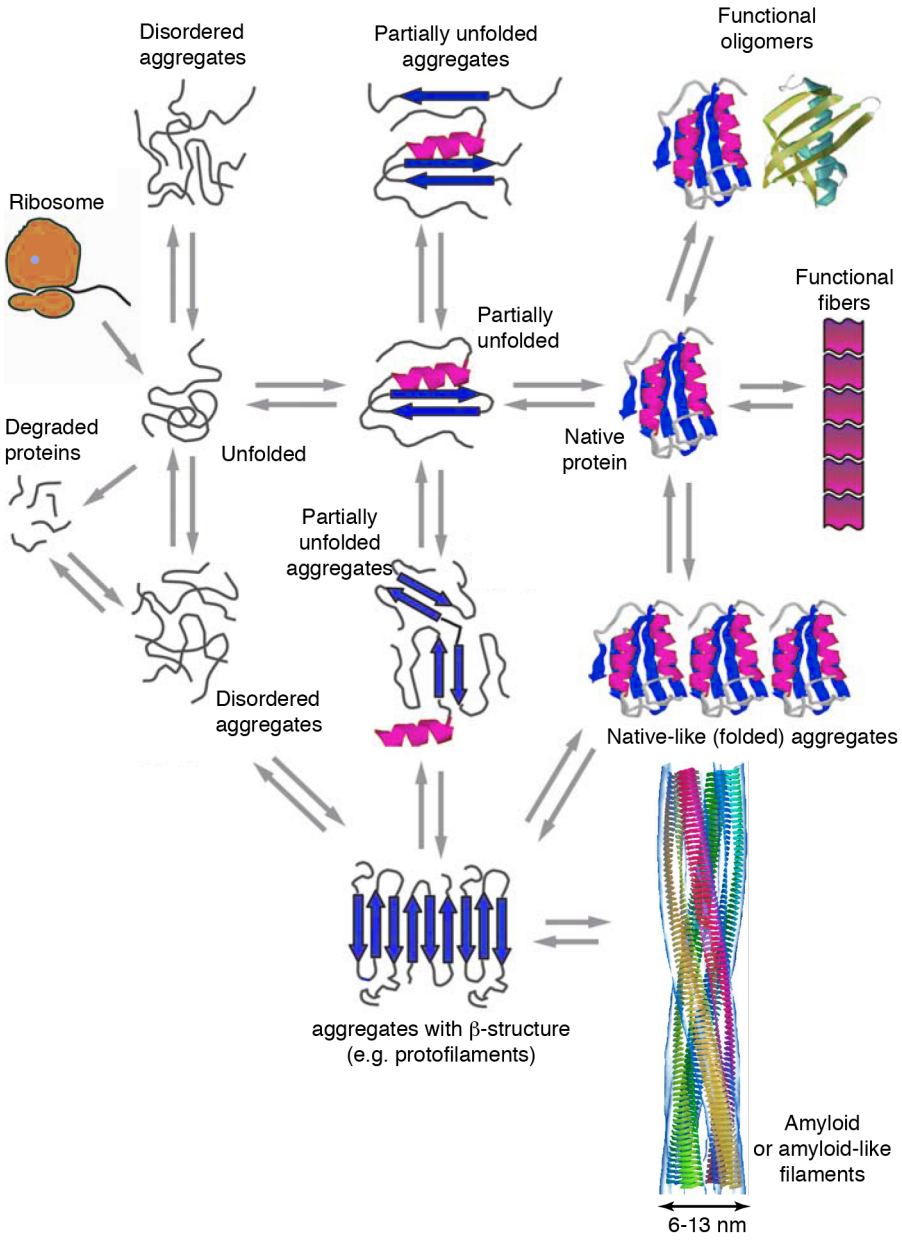

Figure 18: Scheme of pathways driving to amyloid formation. This representation shows the different conformational states available to a polypeptide chain after being translated by the ribosome and the typical pathways of interconversion between them. Such transitions are usually tightly regulated by cells to drive polypeptide chains towards the functional conformation and to avoid aggregation and malfunctioning. Still, for some specific conditions the formation of aggregated species prevails (see section 4.1.1.1). These species can be on-pathway towards amyloid fibrils specially if β-sheet structures are stabilized (bottom region), or off-pathway (top region). Adapted from [117].

well as spherical beads of 2-5 nm in diameter, or structures that resemble beads on a string [139, 145, 146]. Similar amorphous aggregates have been observed for other proteins related to amyloid diseases as for example α-synuclein, polyglutamine proteins, or acylphosphatase. Although this is still a topic of intense research, some of the latest evidences suggest that their mechanism of cytotoxicity could be due to the fact that these unstructured oligomers expose structurally disorganized hydrophobic residues that are able to bind to the plasma membrane causing its disruption with the subsequent effects on cellular homeostasis [147, 148]. Hence there is also an increasing interest in understanding the basic mechanism by which these initially amorphous oligomers self-assemble and interact with other cellular components (e.g. plasma membrane, polysaccharides, nucleic acids).

4.1.2 *Amorphous aggregate formation*

Research on the formation of amorphous aggregates has been largely overshadowed by the more well studied case of structured aggregates such as amyloid. This is likely due to the implication of the later on neurodegenerative disorders that represent a major health issue in aging societies. Moreover, the higher heterogeneity and structural diversity of amorphous aggregates makes its study complex and technically challenging and has precluded research in the field [122, 149]. However, the growing recognition that amorphous oligomers might be the cytotoxic species of amyloid diseases is encouraging research on their structures and mechanisms of interaction [132, 147].

Work in this field is just emerging, but recent efforts point towards the direction that kinetics of amorphous aggregate formation might be fundamentally different from the mechanisms of fibril growth. Amorphous aggregation can not be generally modeled with nucleated growth models characteristic of crystallization or polymerization reactions [120, 150]. However, interpretations based on an analogy to a glass transition, in which concepts such as frustration an supersaturation are incorporated, have been successful in rationalizing experimental data on amorphous aggregation [120]. These works suggest that whereas the kinetics of amyloid formation are slow due to the need to jump over a high energy barrier (i.e. formation of nuclei), amorphous aggregation happens straightforwardly without crossing a high kinetic barrier. In this "glassy dynamics" region, interactions occur on a promiscuous and noncooperative fashion and kinetics are controlled by long-lived low-energy traps, in analogy to previous protein folding models [120, 151]. Indeed this interpretation might also contain a mechanistic basis for nucleation of amyloid fibrils. It is possible that the formation of local and transient glassy states in which an ensemble of amorphous nuclei are formed, could provide the initial amyloid-competent seeds required for amyloid fibrillation.

Other recent approaches based on phenomenological models [150, 152], Ostwald ripening [153], and minimal models of kinetic reactions [154] have also succeeded in providing a theoretical framework and mechanistic insight onto the underlying physics of amorphous aggregation. It is noteworthy that a hallmark of amorphous aggregation is the fact that the rate of aggregation is proportional to the surface area of the particles, in opposition to linear polymerization models [150, 154]. This has been previously observed in nano cluster formation [155], and found as a characteristic result of amorphous aggregation in several experimental and theoretical studies. Finally, different attempts to rationalize non-specific inhibition by unstructured aggregates of organic compounds has also been successful on the pharmacological industry [156, 157].

4.2 THE HYDROPHOBIC EFFECT

A main force driving the formation of both amorphous and structured aggregates are hydrophobic interactions. Here, we briefly review the mechanism by which nonpolar molecules tend to aggregate into higher order structures in water; i.e. the so-called «hydrophobic effect». This effect is of extreme importance in biochemistry and physical chemistry, as it is the main force driving a large number of non-covalent molecular interactions such as: membrane formation, protein folding, self-assembly of amphyphilic micelles, or the insertion of transmembrane proteins into lipid bilayers. The key role of such forces in protein science was put forward by Kauzmann in the 1950's [158] and generalized by Tanford influential papers [159, 160, 161].

Mixing between two substances A and B (other than water), is typically favored by the increase in translational entropy of the mixed ensemble ΔS_{mix} in relation to the two pure separated phases (analogously to an ideal gas):

$$\Delta G_{mix} = \Delta H_{mix} - T\Delta S_{mix} \tag{19}$$

In this general case, insolubility takes place when enthalpic interactions between particles of each of the two substances (i.e. A–A and B–B interactions) are much larger than the gain in translational entropy, making the mixing of the two species unfavorable [162]. However the self-assembly and aggregation observed when nonpolar molecules are placed in water happens due to a very different mechanism. Water molecules have a large dipole moment due to the high electronegativity of the oxygen atom, that has a lone pair of electrons able to electrostatically interact with hydrogen atoms of other water molecules. This creates a highly directional bond known as H-bond (Figure 19**a**). The tetrahedral structure of water allows each water molecule to create several H-bonds simultaneously. In ice for instance, each oxygen

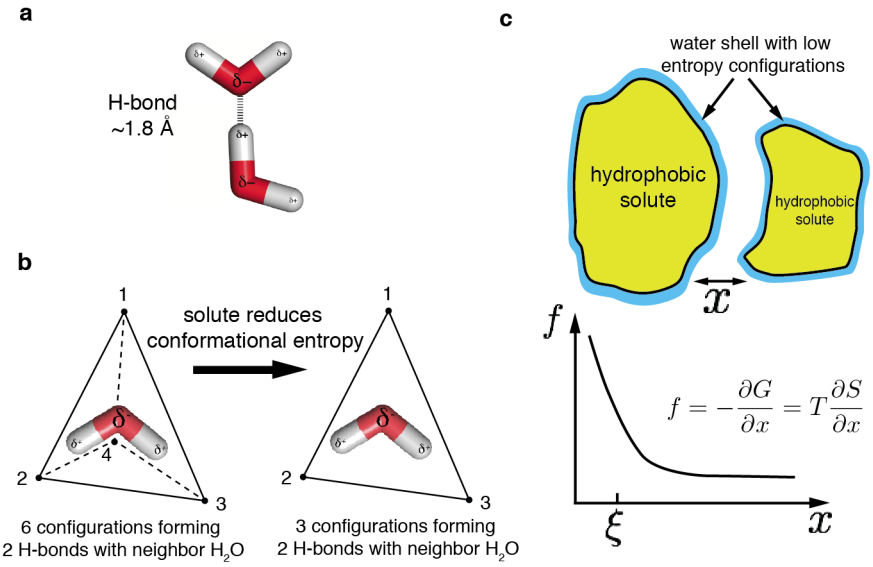

Figure 19: Hydrophobic effect.(**a**) Two water molecules forming an H-bond. The high electronegativity of the oxygen atom (red) causes a dipole moment in water molecules. H-bonds are created due to strong dipole-dipole interactions between oxygen atoms (red) and hydrogen atoms (white). (**b**) Water molecules form a tetrahedral structure in which they are able to form H-bonds with two neighbors (black dots) simultaneously. In the shown example the central water molecule creates H-bonds with neighbors 2 and 3. Hence, in bulk water each molecule can explore 6 H-bonding configurations with different neighbors (1-2, 1-3, 1-4, 2-3, 3-4, 2-4). An hydrophobic solute replaces neighboring water molecules reducing the number of possible H-bonding configurations. (**c**) The reduction of conformational entropy of water molecules around a nonpolar compound, causes an effective attractive force between hydrophobic molecules that tends to reduce the interfacial area between solute and water. This force takes place at distances comparable to a few H-bonded water molecules ($\xi \sim 1$ nm).

atom creates 2 H-bonds with its neighboring molecules in a highly ordered structure in which each water molecule is H-bonding 4 neighbors. In liquid water, the long range order of the crystalline structure is lost in favor of a dynamic network of H-bonds in which in average each water molecule creates still maintains 3.5 H-bonds with its neighbors, showing that local order is preserved. This short-range H-bond network explains for instance the abnormal high boiling point of water if compared to similar molecules (e.g. methanol).

When a nonpolar molecule is placed in this isotropic network of H-bonds, water molecules surrounding the solute have a lower number of possible H-bonding configurations with its neighboring molecules, and thus a reduced conformational entropy (Figure 19**b**). This anisotropy introduced by nonpolar molecules creates an ordering of water molecules around the solute (usually referred to as the clathrate-cage or "iceberg" model [163]), that strongly reduces the entropy of the solution unfavouring mixing. When many such nonpolar molecules are present in a solution, this decrease in entropy creates an effective at-

tractive force between nonpolar molecules that bind together in order to minimize the contact area between solute and water (Figure 19**c**). What is remarkable of the hydrophobic effect therefore is precisely that the mixing of nonpolar molecules into water is not opposed by a large increase in enthalpy ΔH_{mix}, but by a large decrease in entropy ΔS_{mix}. This effective force is known as the «hydrophobic effect» and constitutes a unique organizational force for biological matter [160].

CHAPTER 5

SCOPE AND OBJECTIVES

This thesis aims to investigate the mechanism of binding of small ligands to nucleic acids using single-molecule techniques. Particularly, the thesis focuses in investigating peptide anticancer agents of natural origin that interact with DNA in different ways. Most of the presented work focuses in two compounds (Thiocoraline and Kahalalide F) that are highly insoluble and tend to aggregate and adsorb into surfaces. Therefore, a main objective of this thesis has been to exploit the potentialities of single-molecule techniques to investigate such low solubility compounds that are otherwise difficult to characterize with bulk techniques: from the thermodynamics and kinetics of binding, to the sequence-selectivity and the kinetics of self-association. This objective has been mostly apporached using DNA micromanipulation experiments with optical tweezers; when needed, results have been also complemented with other techniques such as magnetic tweezers, atomic force microscopy or dynamic light scattering. To reach this broad methodological goal, we have focused in accomplishing a series of objectives:

- **Elastic properties of bound dsDNA.** Characterize the elastic response of dsDNA, and determine how the binding of small ligands modifies these elastic properties. Rationalize the observed effects on the basis of the binding mechanism of each ligand, and obtain quantitative information of its equilibrium binding properties.

- **Elastic properties of bound ssDNA.** Develop tools to investigate the elastic response of ssDNA under varying ionic conditions, as well as the formation of non-native secondary structures. Use this methodology to investigate the binding of ligands to ssDNA.

- **Binding kinetics protocols.** Investigate the kinetics of binding of small ligands to nucleic acids using optical tweezers: develop a microfluidics set up and an experimental methodology to obtain reproducible kinetic traces.

- **DNA unzipping assays.** Investigate the potentialities of DNA unzipping experiments to probe DNA-ligand interactions. DNA unzipping provides an alternative reaction coordinate to classical DNA stretching assays, and so such experiments might complement the information obtained from DNA elasticity measurements.

- **Determination of binding pathways.** Development of theoretical models to rationalize the results obtained from kinetics experiments and thermodynamic measurements. Such models can be used with two purposes: obtain a quantitative insight on the thermodynamics and kinetics of the binding reaction, or to test the validity of an underlying microscopic mechanism that explains the macroscopis results.

- **Molecular constructs synthesis.** Set up a consistent methodology to synthesize the different molecular constructs required to perform the previous experiments: DNA stretching constructs, DNA hairpins of varying size and sequence and handle length, as well as DNA-RNA hybrid constructs.

Part II

THERMODYNAMICS AND KINETICS OF THIOCORALINE INTERCALATION TO DNA

Alice: "How long is forever?"
White Rabbit: "Sometimes, just one second."

Lewis Carrol (Alice in Wonderland)

CHAPTER 6

INTRODUCTION

6.1 MOTIVATION

Small DNA binders have become powerful tools in biomedicine due to their ability to interact with nucleic acids at specific sites and modulate key processes such as repair, transcription and replication [31, 114, 164, 165]. Among them, DNA intercalators and bis-intercalators (see section 3) have received continued interest during the last decades due to their particular properties that range from being excellent DNA staining proves (e.g. ethidium, YOYO) to effective anti-proliferative drugs (e.g. doxorubicin, echinomycin) [76, 114]. A detailed explanation of the structure and binding mechanisms of DNA intercalators and bis-intercalators is found in section 3.

Although intercalating agents are widely used in anticancer chemotherapy [65], their mechanism of action based on interfering with the transcription and repair machinery makes them also highly toxic causing undesired side effects [112]. In order to improve their effectivity a lot of effort has been put in developing new ligands that (i) exhibit improved sequence selectivity and (ii) form long-lived complexes [97, 166, 167, 168]. Increased specificity enables recognition and targeting of particular disease-related motifs that can be used for biomedical diagnostics and therapeutics [114, 165]. Slow off-kinetics of long-lived complexes are related to sustained response of anticancer drugs [97, 169, 170]. In addition such ligands may also be useful to develop better staining proves with increased imaging times [86]. Hence, the development of rapid and accurate methods to elucidate the kinetics

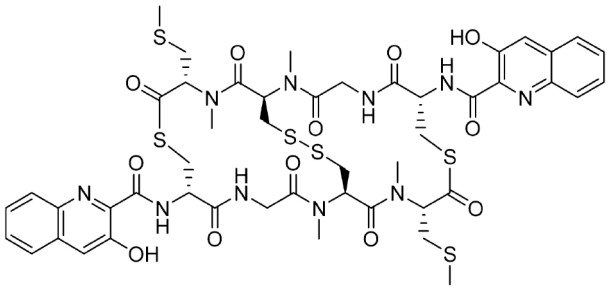

Figure 20: Thiocoraline structure.

and selectivity of small ligands is paramount to advance in the rational design of less toxic and more effective compounds.

Single-molecule force-spectroscopy techniques have proven especially well-suited to investigate the structural and mechanical perturbations that intercalators induce to the DNA double helix [67, 171], succeeding in characterizing the binding of very weak ligands [67, 172]. This is due to the fact that force-spectroscopy uses molecular extension as the natural reaction coordinate and allows to apply force to shift the chemical equilibrium of binding. So far, most experiments have focused on deriving equilibrium information of binding (e.g. affinity) [67, 87, 172, 171, 173], whereas only recently attention has been paid to non-equilibrium experiments from which kinetic rates might be inferred [174, 175, 176]. Moreover, despite the ubiquitous use of DNA intercalators in single-molecule experiments, there is still controversy in some fundamental aspects such as the effects of intercalation on the elasticity of dsDNA [87, 171, 176, 177, 178] and the existence of binding kinetic intermediates [175, 178]. Similarly, other key parameters such as sequence specificity have been scarcely addressed with these techniques, and still rely on classical footprinting assays that are time-consuming and complex to set up and interpret [179, 180].

In this chapter of the thesis, we use optical and magnetic tweezers to study the anticancer bis-intercalator Thiocoraline (Figure 20) [181], fully characterizing the thermodynamics, kinetics and selectivity of binding using a combination of DNA stretching and unzipping assays [182]. A clear understanding of these properties was lacking due to the low solubility of this drug [183, 184], making measurements difficult and leading to limited and conflicting results [185, 186]. We find that Thiocoraline bis-intercalates DNA with an extremely slow dissociation constant that, interestingly, decreases steeply with force. This makes possible to measure the elastic properties of intercalated DNA by performing non-equilibrium pulling experiments in which the binding fraction of ligand remains quenched. We show that bis-intercalation does not modify DNA persistence length whereas the stretch modulus does increase with the bound fraction. A series of

kinetic experiments demonstrate the existence of a mono-intercalated intermediate state that contributes to the long lifetime of the complex. Finally, we develop a single-molecule footprinting assay based on DNA-unzipping, in which we show that at low concentrations the peptide binds DNA in a sequence-specific manner, with a preference to clamp CpG steps flanked by A-T base pairs. Such information is difficult to obtain in bulk assays and cannot be accessed in usual single-molecule stretching assays.

We propose that this sequence-selectivity together with the force-dependent off-rate are an important feature of the mechanism of action of this subnanomolar drug ($IC_{50} \sim 200$ pM). On a broader perspective, we provide new insights on the general mechanisms by which bis-intercalators bind DNA and their kinetic stability. The proposed single-molecule approach will be of great utility to quantify essential parameters of the specific and nonspecific binding modes of ligands that target DNA [76, 179, 187].

CHAPTER 7

RESULTS: PULLING EXPERIMENTS OF THIOCORALINE BIS-INTERCALATING DNA

7.1 FORCE-EXTENSION CURVES OF BIS-INTERCALATION

We pulled single dsDNA molecules (half λ-DNA, 24 kb, synthesis details in Appendix D) in the presence of Thiocoraline in the optical tweezers setup (Figure 21**a**, Inset). DNA is tethered between two beads and stretched at a low force (2 pN) until it reaches its equilibrium extension. The molecule is then repeatedly pulled by displacing one bead relatively to the other, and the extension measured as a function of the applied force. In this way force-extension curves (FECs) are recorded in the range 0-40 pN.

At a concentration of 100 nM, the intercalated DNA has a molecular extension significantly longer (~ 30%) than that of naked dsDNA (Figure 21**a**, red and black respectively), as expected from an intercalative binding mode. Notably, DNA stretching experiments performed at varying pulling speeds show a hysteresis that increases with decreasing pulling speed, whereas hysteresis is not observed at the fastest speed of 1.5 μm/s (Figure 21**a**, red, blue, green). The observation of hysteresis in DNA pulling curves was previously reported for the bis-intercalator dye YOYO-1 [178], and is caused by the fact that the affinity constant and kinetic rates are force-dependent. It has been shown that the affinity constant of DNA intercalators increases exponentially with applied force [67], and therefore at equilibrium more ligands are bound at higher forces. In a typical pulling cycle the force is increased from a low value (2 pN) up to a high value (40 pN) and

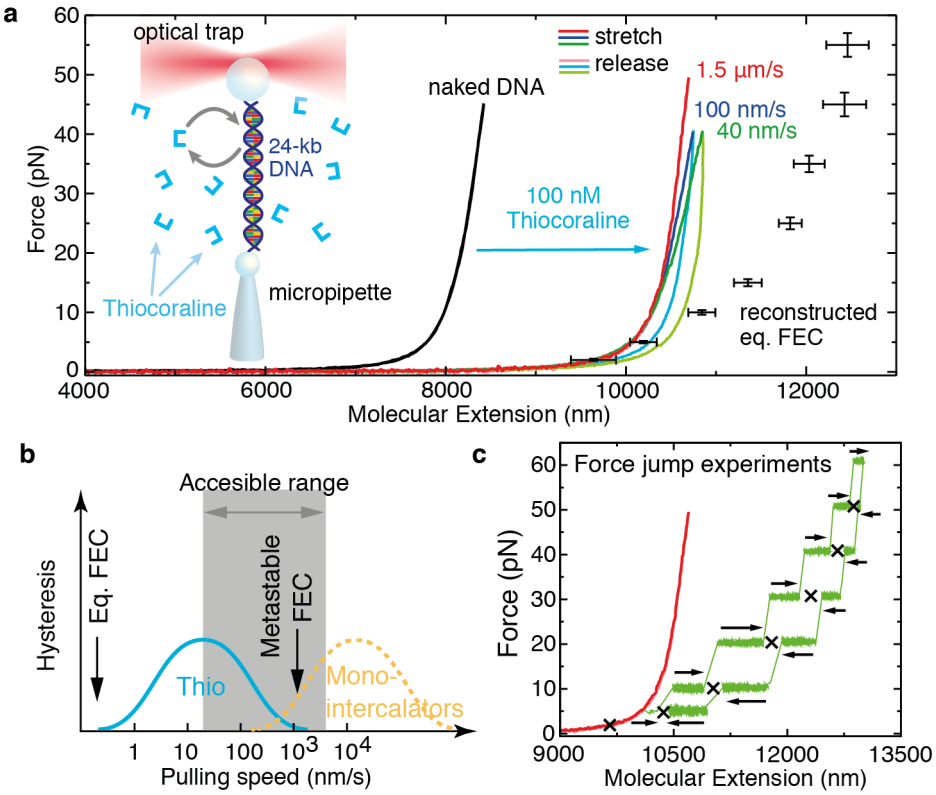

Figure 21: FECs of bis-intercalated DNA. (a) DNA pulling curves before (black) and after flowing 100 nM Thiocoraline at varying pulling speeds: 1.5 μm/s (red), 100 nm/s (blue), 40 nm/s (green). Dark (light) colors correspond to the stretching (releasing) parts of the cycle. The equilibrium FEC (black points, mean± SD, N⩾5) is recovered from force jump experiments. **(Inset)** Scheme of the DNA stretching setup. **(b)** In force-jump experiments the relaxation of the molecular extension (green) is measured at constant force values starting from different initial conditions in bidirectional experiments (arrows). The equilibrium extension at each force (crosses) can then be inferred. A metastable FEC (1.5 μm/s) is shown in red as a reference. **(c)** Illustrative scheme of the hysteresis observed in FECs as a function of pulling speed. Thiocoraline (cyan) shows an optimal range to measure metastable FECs, as compared to mono-intercalators (yellow) that reach equilibrium over shorter timescales.

then decreased again, enhancing intercalation during a pulling cycle. Two are the relevant timescales in a pulling experiment: the equilibration time of the binding reaction (τ_{eq}) and the observational time (the duration of a pulling cycle, τ_{obs}). If pulls were done slowly enough ($\tau_{obs} > \tau_{eq}$), the equilibrium FEC would be obtained. In this case we would expect the fraction of bound ligands to increase with force. In the opposite regime ($\tau_{obs} \leqslant \tau_{eq}$), the molecular extension cannot relax to its equilibrium value at each force, and a positive hysteresis is observed (blue and green FECs in Figure 21**a**). Remarkably enough, if the pulling speed is increased ($v > 1$ μm/s, red) metastable FECs are measured, where non-equilibrium but reversible curves are obtained. Along these metastable FECs the number of bound intercalators remains quenched, and the FEC does not show hysteresis (Figure 21**b**). To the best of our knowledge, this metastable state was not attainable

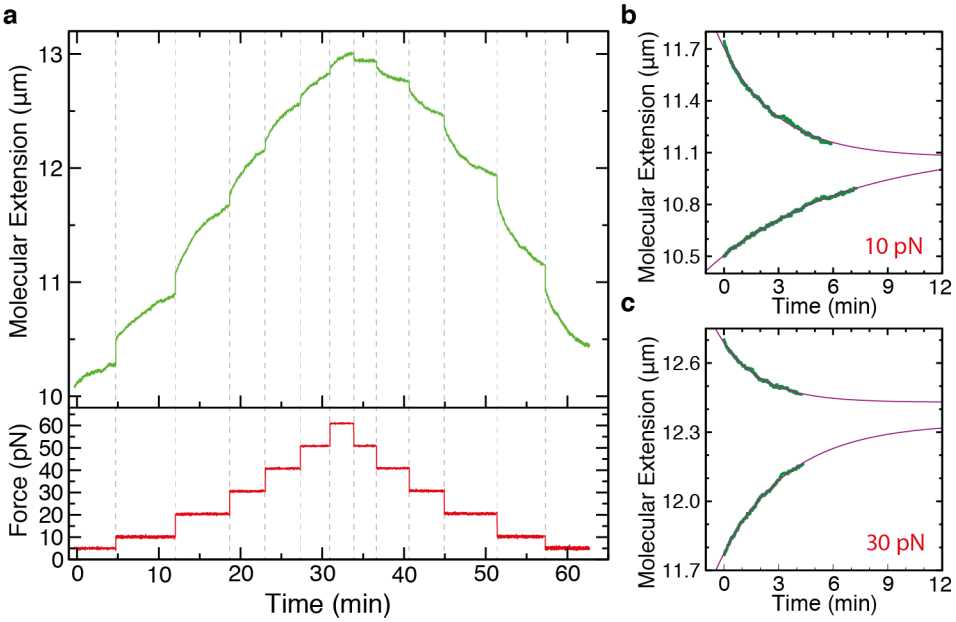

Figure 22: Typical bidirectional force-jump experiment. The force (red) is increased stepwise up to a maximum value using force-feedback and then decreased down to the initial value using a reversed protocol. In this way, the relaxation of the molecular extension (green) is obtained in the forward and backward direction for different force values. **(b)** Relaxation of the molecular extension (green) for two forces from panel (a). Data is fitted to a single exponential function (purple) from which the equilibrium extension $x_{eq}^{[Thio]}(F)$ at each force is determined. Experiments correspond to 100 nM Thiocoraline.

with previously studied mono-intercalators and bis-intercalators due to their faster kinetic rates. In section 7.2, we will exploit this feature of metastability to determine the elastic properties of bis-intercalated DNA for a fixed fraction of ligand.

On the other hand, the slow kinetics of Thiocoraline does not allow us to obtain an equilibrium FEC using a standard pulling protocol (Figure 21**b**). To do so, we performed force-jump experiments in the range 5-60 pN (Figure 21**c**). In these experiments, the molecule is maintained at a constant force using force-feeedback, and the relaxation of the molecular extension is measured at each force in bidirectional experiments (black arrows in Figure 21**c**). As shown below (section 7.1.1), relaxation curves of force-jump experiments are well-described by a single-exponential function, and their asymptotic value can be used to determine the equilibrium extension at each force (Figure 21**a**, black points). A comparison between the reconstructed equilibrium FEC and the metastable FEC (Figure 21**a**, red curve and black points respectively), illustrates the pronounced effect of force in shifting the binding equilibrium of Thiocoraline.

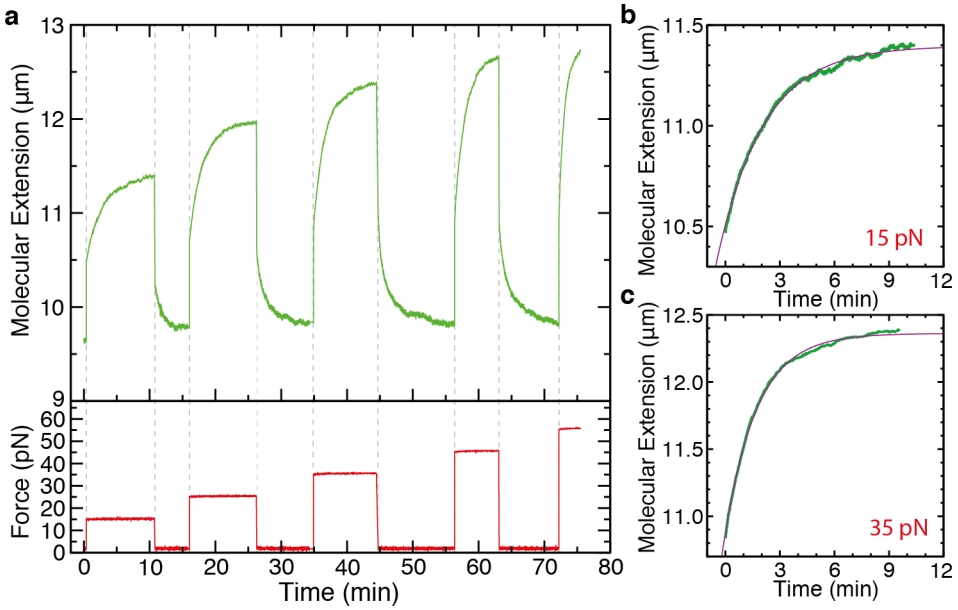

Figure 23: Typical unidirectional force-jump experiment. (a) The force (red) is increased from 2 pN to a preset value, and the relaxation of the molecular extension (green) is measured until reaching saturation. Then the force is decreased again to 2 pN and the measurement repeated for a set of different forces. **(b)** Relaxation of the molecular extension (green) at two forces from panel (c). Data is fitted to a single exponential function (purple), obtaining $x_{eq}^{[Thio]}(F)$ and the total kinetic rate k_{tot} at each force. Experiments correspond to 100 nM Thiocoraline.

7.1.1 *Force-jump experiments*

The experiments were performed in two ways: (i) bidirectional stepwise force-jumps, and (ii) unidirectional force-jumps from low force. In the first type of experiments (Figure 22**a**), the force was increased stepwise every ~5 min up to a maximum preset force. Then the protocol was reversed and the force stepwise decreased down to the initial value. By doing so, we obtained a forward and backward relaxation curve for each force. The relaxations were well described by a single-exponential function, from which we obtained the forward and backward asymptotic extensions (Figure 22**b,c**). The equilibrium extension at each force $x_{eq}^{[Thio]}(F)$ was estimated as the average of both values. To show that $x_{eq}^{[Thio]}(F)$ does not depend on the stretching history, we independently measured this value from unidirectional force jumps (Figure 23**a**). In these experiments, the force was increased from 2 pN to a preset value, and the relaxation of the molecular extension was followed until reaching saturation (~ 10 min). Consecutive force-jumps between 2 pN and the different preset forces were performed in this way, and fitted to a single-exponential function (Figure 23**b,c**). The $x_{eq}^{[Thio]}(F)$ values obtained with both methods were fully compatible. Moreover, from the unidirectional force-jump experiments we also obtained the total kinetic rate $k_{tot}(F)$ at each force (see section

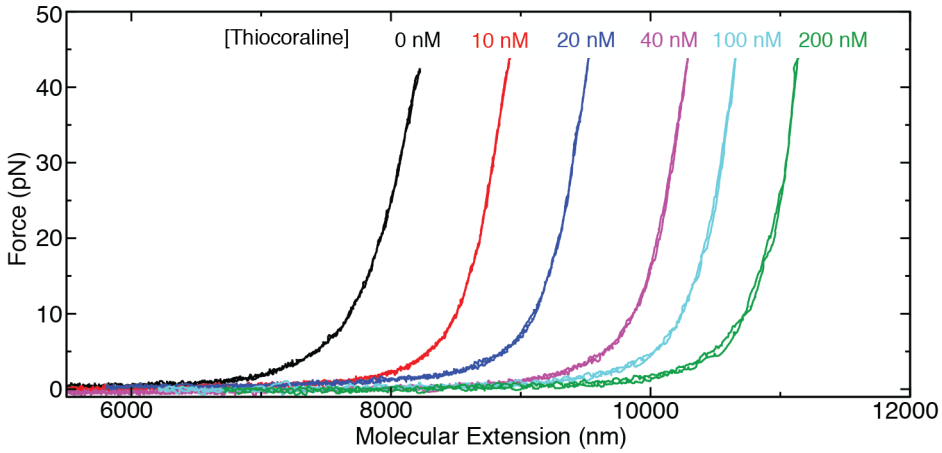

Figure 24: Metastable FECs. Metastable FECs at different Thiocoraline concentrations: 10 nM (red), 20 nM (blue), 40 nM (purple), 100 nM (cyan) and 200 nM (green). A FEC of naked dsDNA is shown in black as a reference.

8.1). All force jumps were performed with a force-feedback operating at 1 kHz in which force was changed in less than 0.1 s and maintained constant to the preset value.

7.2 THERMODYNAMICS OF BINDING AND ELASTIC PROPERTIES OF INTERCALATED DNA

7.2.1 *Metastable FECs and analysis with elastic polymer models*

The slow kinetics of Thiocoraline makes possible to determine the elastic properties of intercalated DNA in conditions of metastability for a fixed fraction of bound ligand. In our experiments, such fraction is initially equilibrated at a force of 2 pN and the molecule is repeatedly pulled up to 40 pN fast enough to have that fraction quenched along the pulling cycle (v=1.5-3 μm/s). We performed pulling experiments over 3 decades of ligand concentration (0.5-1000 nM, Figure 24), and fitted the metastable FECs to the extensible worm-like chain (WLC) model (Equation 56). For each tethered molecule we determined the l_0, L_p and S values by fitting a set of consecutive pulling cycles (typically more than 5). The reported values at each concentration is the average of a set of different molecules ($N \geqslant 7$).

The contour length increased with Thiocoraline concentration (Figure 25**a**) reflecting the varying fraction of bound intercalator [63]. As explained below (section 7.2.2), these results can be fitted to the McGhee-Von Hippel (MGVH) model (Figure 25**a**, cyan), from which characteristic equilibrium values of the interaction (the binding affinity K_d and binding-site size n) are obtained. As explained in the following chapter (section 8.3) we also developed a three-state kinetic model that explains the kinetics of binding of the ligand. Remarkably enough, the equilibrium elongations predicted by this model at dif-

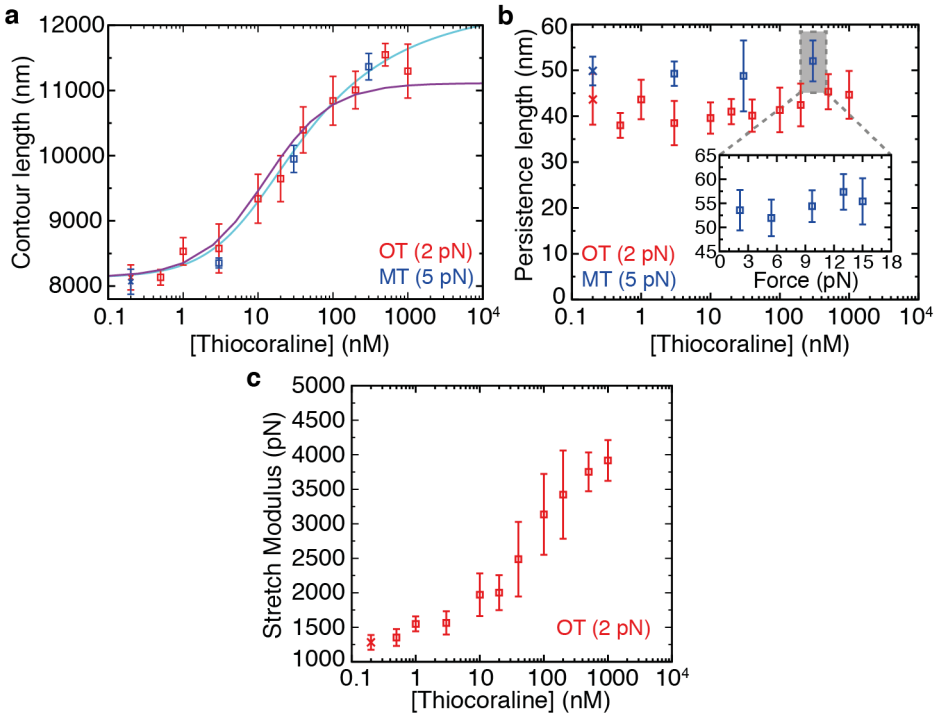

Figure 25: Elastic properties of bis-intercalated DNA. (a) Contour length (l_0) of DNA as a function of [Thiocoraline] in optical (red squares) and magnetic tweezers (blue squares) experiments. The optical tweezers data is fitted to a MGVH binding isotherm (cyan, Eq. 21) to obtain the equilibrium affinity and binding-site size at low force ($K_d = 77 \pm 13$ nM, n=3.84 $\pm$ 0.12 bp). A fit to a three-state model as explained in section 14.4.1 is shown in purple. **(b)** Persistence length (L_p) of DNA as a function of [Thiocoraline]. L_p remains independent of [Thiocoraline] both in optical tweezers (red) and magnetic tweezers (blue) experiments. **(Inset)** L_p remains also unaffected when the bound fraction is changed by equilibrating the molecule at a different force. **(c)** Stretch modulus (S) of DNA as function of [Thiocoraline]. S monotonically increases with [Thiocoraline]. Values are reported as mean$\pm$SD (N$\geqslant$7), and are determined by fitting metastable FECs (v=1.5-3μm/s) to the extensible (optical tweezers) or inextensible (magnetic tweezers) WLC model. Reference values from experiments without Thiocoraline for each technique are shown at the lowest concentration (red and blue crosses).

ferent ligand concentrations agree very well with these experimental measurements (Figure 25**a**, purple).

Fits to the WLC model also yielded the persistence length (L_p) and stretch modulus (S) of a bis-intercalated DNA molecule. The persistence length remained constant within experimental errors over the whole range of concentrations investigated (Figure 25**b**). This is in contrast with previous studies for other mono and bis-intercalators in which a strong decrease of the persistence length with ligand concentration was reported [67, 171, 177, 178, 188]. In these works, equilibrium FECs were used to obtain the elastic properties of intercalated DNA. However, in the equilibrium ensemble, the fraction of intercalated drug changes as force is increased, leading to a continuously changing contour length within a pulling cycle, thereby giv-

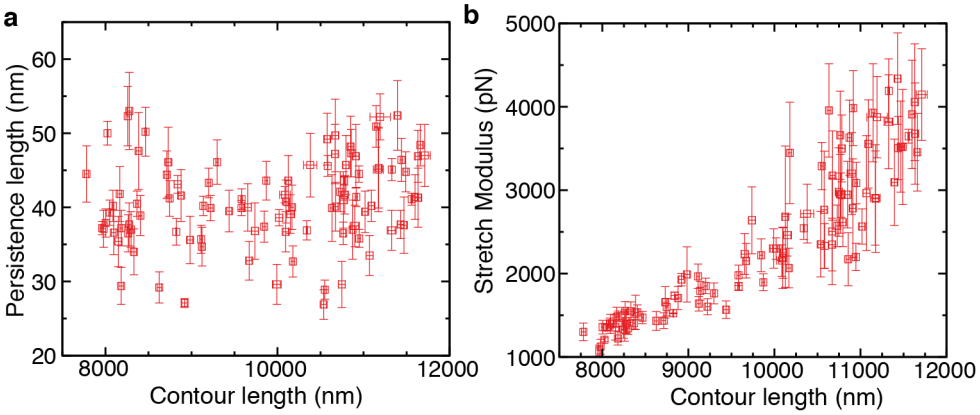

Figure 26: Elastic parameters for individual molecules. (**a**) Persistence length (L_p) as a function of contour length. L_p is independent of the contour length of the bis-intercalated DNA molecule. (**b**) Stretch modulus (S) as a function of contour length. S monotonically increases with contour length, demonstrating that S is directly related to the fraction of bis-intercalator bound to the DNA molecule. A set of consecutive pulling cycles for each molecule are used ($N > 5$), and results are shown as mean± SD. Experiments are performed at varying concentrations between 0.5-1000 nM Thiocoraline.

ing lower effective persistence lengths in the fits. A constant persistence length has also been reported for YOYO-1 in recent studies performed at low forces ($F < 10$ pN) and low salt condition (10 mM) where force-dependent intercalation is negligible [87]. We confirmed our observation by performing experiments with a magnetic tweezers setup in the low-force regime (0.01-10 pN), obtaining equivalent results (Figure 25**b**, blue). Furthermore, we performed experiments in which the fraction of bound intercalator was increased by equilibrating the molecule at a higher initial force (Figure 25**b**, Inset), finding again that persistence length is independent of the fraction of intercalated drug regardless of the method used to increase the binding ratio (i.e. ligand concentration or force). Finally, from our high-force ($F > 10$ pN) measurements we also obtained the stretch modulus of bis-intercalated DNA, finding a systematic increase in the stretching rigidity with ligand concentration (Figure 25**c** and Figure 26).

7.2.2 *Analysis of equilibrium parameters using the MGVH model*

The elongation of the DNA molecule at each Thiocoraline concentration ($l_0^{[Thio]} - l_0^{DNA}$) is directly proportional to the number of intercalators bound to the molecule, and to its equilibrium binding fraction (ν):

$$\nu = \frac{\#\text{ bound ligands}}{\#\text{ of DNA bp}} = \frac{l_0^{[Thio]} - l_0^{DNA}}{\Delta l_0 N_{bp}} = \frac{l_0^{[Thio]} - l_0^{DNA}}{2 l_0^{DNA}} \qquad (20)$$

where $l_0^{DNA} = N_{bp} d$ (d being the base-pair rise, 0.34 nm) and Δl_0 the DNA elongation induced by one bis-intercalator. We have

assumed that each bis-intercalator elongates the DNA molecule a distance twice d ($\Delta l_0 = 2d$) [63, 67]. For ligands that bind non-specifically and non-cooperatively to a one-dimensional lattice such as DNA, the MGVH model [189] predicts that the equilibrium binding fraction satisfies the implicit equation:

$$\nu = \frac{[L]}{K_d} \frac{(1 - n\nu)^n}{(1 - n\nu + \nu)^{n-1}} \tag{21}$$

where K_d is the ligand affinity constant, n is the binding-site size (number of bp covered by each ligand), and $[L]$ the ligand concentration. By combining Eqs. 20 and 21 we can fit the titration results (Figure 21**c**), and obtain the equilibrium affinity constant $K_d = 77 \pm 13$ nM and binding-site size $= 3.84 \pm 0.12$ bp of Thiocoraline (compatible with 1 ligand every 4 base pairs). The latter is in good agreement with the values previously determined for other DNA bis-intercalators and from X-ray crystallography [77, 79, 173]. A Levenberg-Marquadt algorithm was used both for the WLC and MGVH fitting routines.

7.3 PULLING EXPERIMENTS IN THE OVERSTRETCHING REGION

The DNA pulling curves presented in the previous sections are performed up to a maximum stretching force of 40 pN. This force is below the overstretching transition (F_{OS} ~63 pN at 100 mM NaCl), and allows us to obtain metastable FECs at the highest pulling speed of 1.5 μm/s for all Thiocoraline concentrations ([Thio]$\leqslant$1 μM). However, if the maximum pulling force is increased to 70 pN, FECs show longer molecular extensions and larger hysteresis as compared to 40 pN, due to the greater intercalation achieved at the highest forces (Figure 27**a**). In fact, at 100 nM hysteresis is observed also at the fastest pulling speed of 1.5 μm/s (Figure 27**b**), and we cannot pull fast enough to reach the metastable regime at this higher force (in contrast to the metastable FECs obtained at this same concentration when the maximum force was 40 pN, Figure 21). As previously observed for other intercalators [178, 190], a tilted plateau related to the overstretching transition is observed. If experiments are performed at lower bis-intercalator concentrations, the equilibration time of the binding reaction τ_{eq} increases and the metastable FECs can be obtained again at the fastest pulling speed (Figure 27**d-f**, red). As expected, when we decrease ligand concentration the molecular extension becomes shorter (due to the lower fraction of ligand bound at low forces), and the slope of the force plateau decreases approaching the overstretching curve of naked dsDNA. These findings are in agreement with a mechanism in which the bis-intercalator binds to B-DNA, elongating and unwinding the double helix, and increasing the energy required to force-melt or fully unwind the bis-intercalated

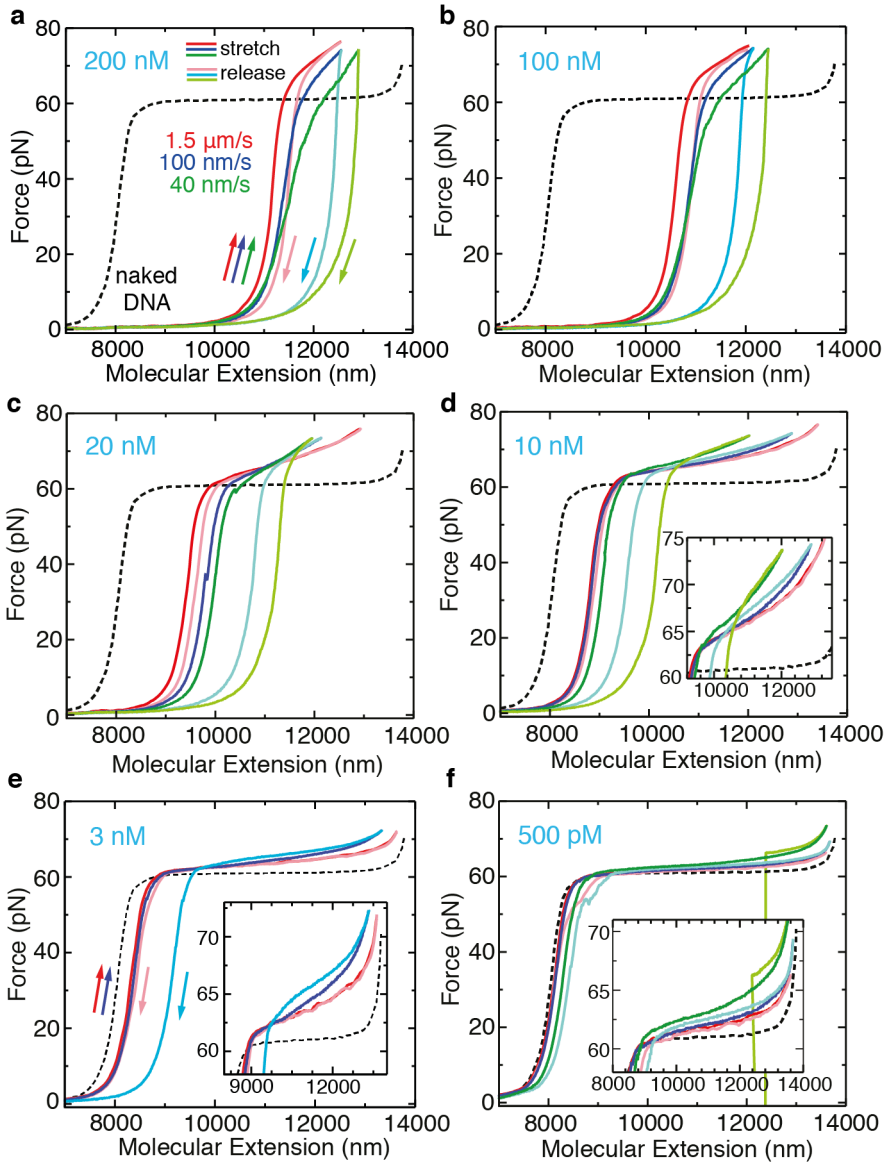

Figure 27: Hysteresis in Thiocoraline pulling experiments. FECs of DNA at six different Thiocoraline concentrations: 200 nM (a), 100 nM (b), 20 nM (c), 10 nM (d), 3 nM (e), 500 pM (f). Experiments are performed up to a maximum force of 70 pN at three different pulling speeds: 1.5 µm/s (red), 100 nm/s (blue) and 40 nm/s (green), and dark and light colors correspond to the stretching and release part of the pulling cycle respectively. A FEC of naked dsDNA is shown as a reference (dotted black).The sudden force drop to 0 pN in the 40 nm/s (green) experiment in panel (f) corresponds to the breakage of the bead-DNA attachment.

DNA (converting B-DNA into S-DNA). At concentrations below 10 nM, fast pulling experiments show a molecular extension slightly longer (<5%) than naked dsDNA and an overstretching plateau that is slightly tilted, indicative of a few bound intercalators (Figure 27**e, f**). Remarkably, at this low concentrations slower pulling rates show negative hysteresis along the overstretching region (the force is higher in the releasing as compared to the stretching part of the pulling cycle, Figures 27**d-f**). These results suggest that Thiocoraline is able to bind either force-melted regions or S-DNA, stabilizing the DNA double helix and therefore reducing the molecular extension at forces above the overstretching transition. In these experiments, once the force is reduced below the overstretching transition, the molecular extension remains longer than during stretching, confirming that Thiocoraline has bis-intercalated DNA at forces above the overstretching plateau. This is particularly visible in the 3 nM experiment (Figure 27**e**). In this condition, the stretch part of the cycle for the fast (dark red) and slow (dark blue) pulling rate are very similar below the overstretching region. However, during releasing at the slow pulling rate (light blue) the apparent molecular extension is ~700 nm longer than that observed at a fast pulling rate (light red), suggesting that most of the binding has happened at $F \geqslant F_{OS}$. Recent evidence that force-melted DNA and S-DNA coexist at physiological conditions, and their similar extension, makes difficult to unambiguously attribute binding to any of both [47, 48].

CHAPTER 8

RESULTS: KINETIC EXPERIMENTS OF THIOCORALINE BIS-INTERCALATION

8.1 TOTAL KINETIC RATE AND DIRECT OFF-RATE MEASUREMENTS

To characterize the remarkably slow kinetics of Thiocoraline we performed force-jump experiments at two different concentrations (Figure 28**a, b**). The relaxation of the molecular extension at each force allowed us to characterize the total kinetic rate of intercalation as a function of force. Relaxation curves are well described by single-exponential kinetics, with a total rate that increases exponentially with force at both concentrations (Figure 28**c**). For a bimolecular reaction where k_{on} is much larger than k_{off}, the total rate ($k_{tot} = k_{on} + k_{off}$) should be proportional to ligand concentration. However, the total kinetic rates observed at 1 µM were found to be only ~2-fold faster than those obtained at 100 nM, whereas at 10 nM we observed a ~10-fold decrease in relation to 100 nM, suggesting that binding is not a purely bimolecular process.

To verify this hypothesis, we directly measured the off-rate $k_{off}(F)$ from wash-off experiments. In these experiments, an intercalated DNA molecule is hold at a constant force in the optical tweezers setup, while the chamber is flushed with peptide-free buffer (Figure 29). However, the low solubility and hydrophobicity of Thiocoraline makes difficult to fully remove the peptide from the chamber in such experiments. In order to create an effective depletion of ligands, we flushed scavenger DNA. Thus, ligands that unbind from the tethered molecule or from the microfluidics chamber are sequestered and

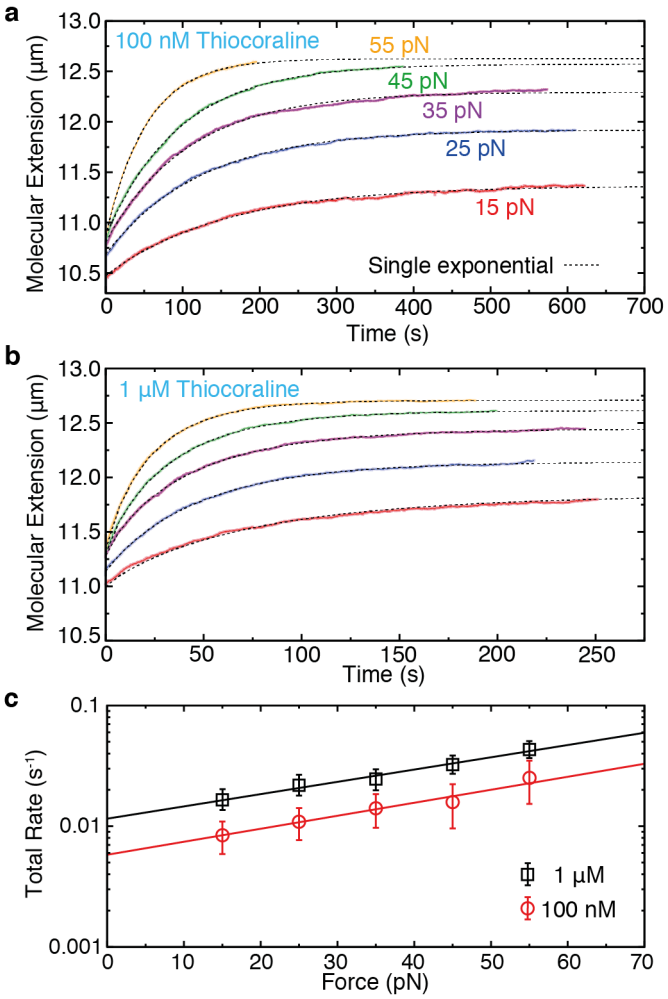

Figure 28: Total macroscopic kinetic rate (k_{tot}) from force-jump experiments. (a, b) DNA extension as a function of time after increasing the force from 2 pN to a preset value of: 15 pN (red), 25 pN (blue), 35 pN (purple), 45 pN (green), 55 pN (yellow). Relaxations are well-described by a single-exponential function and allow us to obtain k_{tot} at the two concentrations investigated: 100 nM in panel (a) and 1 µM in panel (b). **(c)** Force-dependence of k_{tot} at the two investigated concentrations. The intercalation-rate increases exponentially with force with similar slopes for both concentrations.

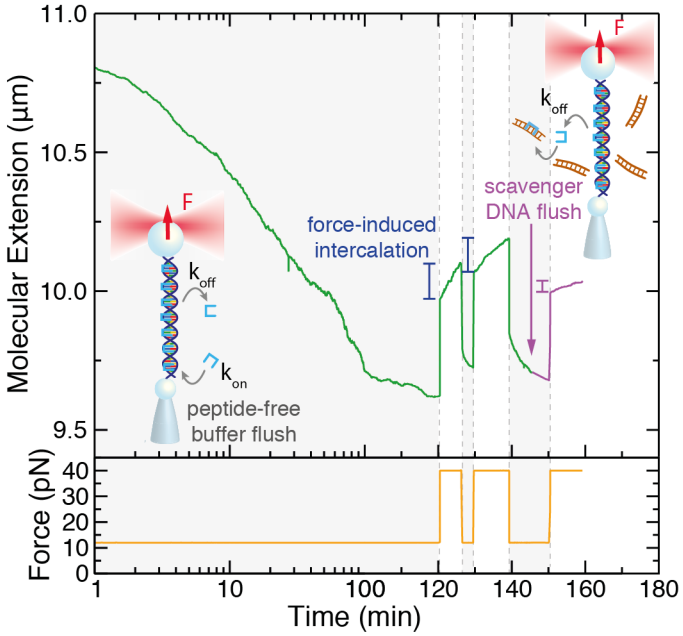

Figure 29: Wash-off experiments at constant force. DNA extension as a function of time (green) in a wash-off experiment performed at 12 pN (orange) using peptide-free buffer. The extension decreases with time due to the unbinding of Thiocoraline. However, upon increasing the force after 2 hours, an exponential elongation is observed, showing that small amounts of ligand remained in the microfluidics chamber. This effect can be suppressed by transiently flushing a high concentration of scavenger DNA (purple).

dragged away. For instance, in the experiment shown in Figure 29, even after 2 hours of peptide-free buffer flush, force-induced intercalation occurs if the force is increased to 40 pN. However, a brief flush of 10 μM bp DNA largely blocks this effect in a subsequent force-jump (Figure 29, purple). This shows that ligand-free buffer flow is not sufficient to wash the chamber in constant-force experiments.

The scavenger DNA solution used for the wash-off experiments was prepared by digestion of λ-DNA with a blunt-ended restriction enzyme (EcoRV) that creates fragments sized 250bp-6kb. In this way, a randomly assorted distribution of sequences equivalent to those in the tethered DNA molecule was obtained. The digested DNA was dissolved to a 10 μM bp concentration in the same buffer than the ligand samples (Tris-EDTA [TE] buffer 100 mM NaCl, pH7.5) and flowed into the experimental area using an auxiliary capillary glass tube. A low flow was maintained during the experiments ($F_{drag} \leqslant 2$ pN) to ensure ligand removal.

8.1.1 *Verifying ligand depletion during wash-off experiments with DNA unzipping*

To test how reliable is this approach to generate ligand depletion, we performed unzipping experiments with a long DNA hairpin (6.8

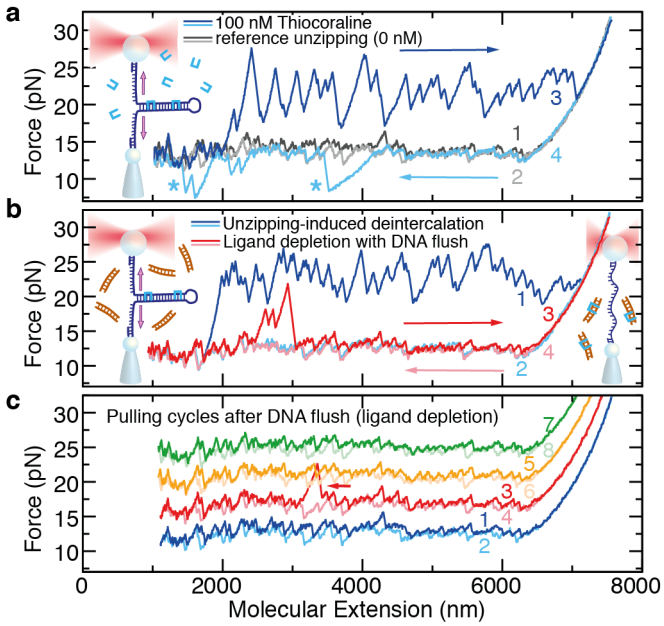

Figure 30: Wash-off experiments using scavenger DNA. (a) Unzipping the DNA hairpin (6.8-kb) without ligand shows a characteristic quasi-reversible sawtooth pattern due to base-pair disruption (gray). In the presence of Thiocoraline, a series of force peaks are observed during pulling (dark blue) indicative of binding events. During refolding (light blue), kinetic states that temporarily block the hairpin rezipping are sometimes observed (*). **(b)** To test the wash-off method, we flushed the scavenger DNA in, and removed the bound ligands by unzipping and rezipping the hairpin (blue). If the hairpin is immediately unzipped again (red), very few binding events remain. **(c)** Consecutive unzipping cycles performed during the following 10 min do not show binding (cycles shifted upwards for clarity), and only occasionally (< 20%) individual binding events are observed (arrow).

kb). In DNA unzipping experiments the hairpin is pulled from the two ends and base pairs are sequentially disrupted (Figure 30**a**). In the absence of ligands, the unzipping shows a characteristic sawtooth pattern at a force ~ 15 pN that is quasi-reversible at the experimental pulling speed (Figure 30**a**, dark gray is unzipping and light gray rezipping). In the presence of 100 nM Thiocoraline (Figure 30**a**, dark blue), the unzipping pattern shows a series of peaks at $F \geqslant 20$ pN due to the increased DNA stability at the positions where bisintercalators are bound. In these experiments, the molecule can be rezipped again by decreasing the distance between beads (Figure 30**a**, light blue), recovering the typical sawtooth pattern that indicates DNA hybridization. Interestingly, in the presence of Thiocoraline, a few specific locations of the hairpin (indicated with an $*$) require lower forces to refold. We attribute this rare events to the formation of competing non-native structures that are locally stabilized by Thiocoraline, and delay the formation of the native hairpin. The molecule can then be repeatedly pulled in the presence of 100 nM Thiocoraline, obtaining equivalent unzipping (dark blue) and rezipping (light blue) curves.

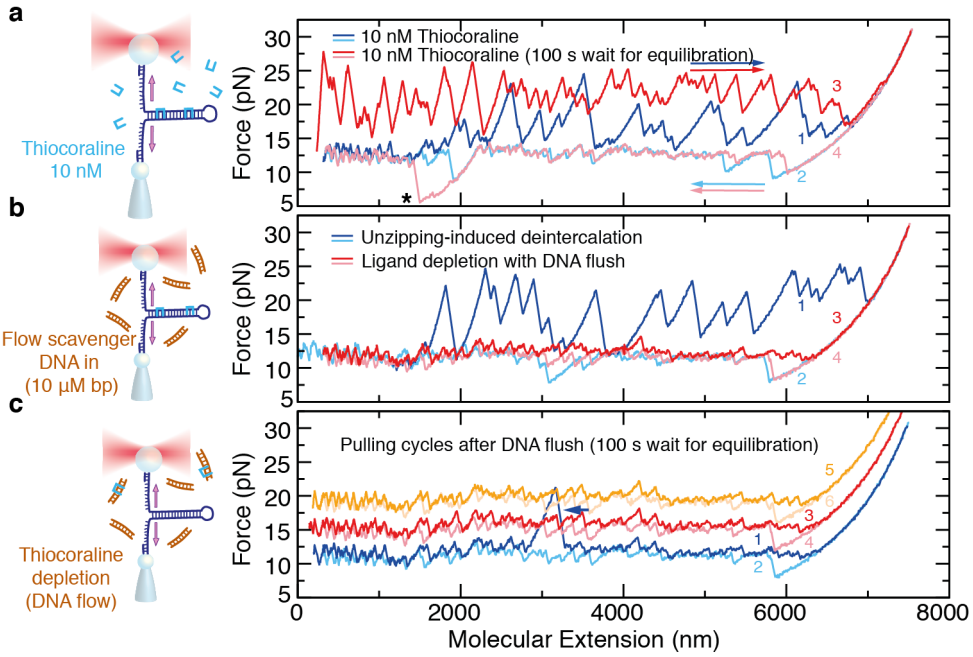

Figure 31: DNA unzipping at 10 nM Thiocoraline and wash-off with scavenger DNA. (**a**) Unzipping of the 6.8-kb DNA hairpin at 10 nM Thiocoraline (dark blue) shows a lower fraction of binding events than at 100 nM (Figure 30**a**, blue) as expected. The number of bound intercalators can be kinetically increased by waiting 100 s between consecutive pulling cycles (dark red). Interestingly, at this much lower concentration the kinetic stabilization of non-native structures at specific locations (*) is also observed in the rezipping curves (light red and blue). (**b**) The bound ligands are removed by unzipping (dark blue) and rezipping (light blue) the hairpin, after flowing the scavenger DNA solution for 3 min. If the hairpin is immediately unzipped again (red), we do not observe any binding event showing that the experimental area is free of ligand. (**c**) Consecutive unzipping cycles performed during the following 15 min are indistinguishable from naked DNA (cycles are shifted upwards for clarity) and only very occasionally individual binding events (blue arrow) are observed. Pulling speed is 500 nm/s.

To test the DNA wash-off method, we flushed the scavenger solution while holding the intercalated DNA molecule partially unzipped. After a 3 min flush, the molecule is unzipped to remove all the intercalated peptide and rezipped again (Figure 30**b**, blue). In the following unzipping cycle, only 3 binding events are detected (Figure 30**b**, red). Successive pulling cycles are indistinguishable from naked DNA unzipping experiments (Figure 30**c**), and only individual binding events (e.g. red FEC in Figure 30**c**) are occasionally observed (<20%). This assay is therefore particularly well suited to detect subnanomolar concentrations of ligands, as even single binding events give a clear footprint in the DNA unzipping pattern (i.e. over 6800 potential binding sites). We conducted similar experiments starting with a lower initial concentration of 10 nM Thicoraline with equivalent results (Figure 31). Consequently, by using scavenger DNA in the wash-off experiments, it is possible to create a depletion of ligand in the experimental area and, as shown below, perform accurate off-rate measurements at dif-

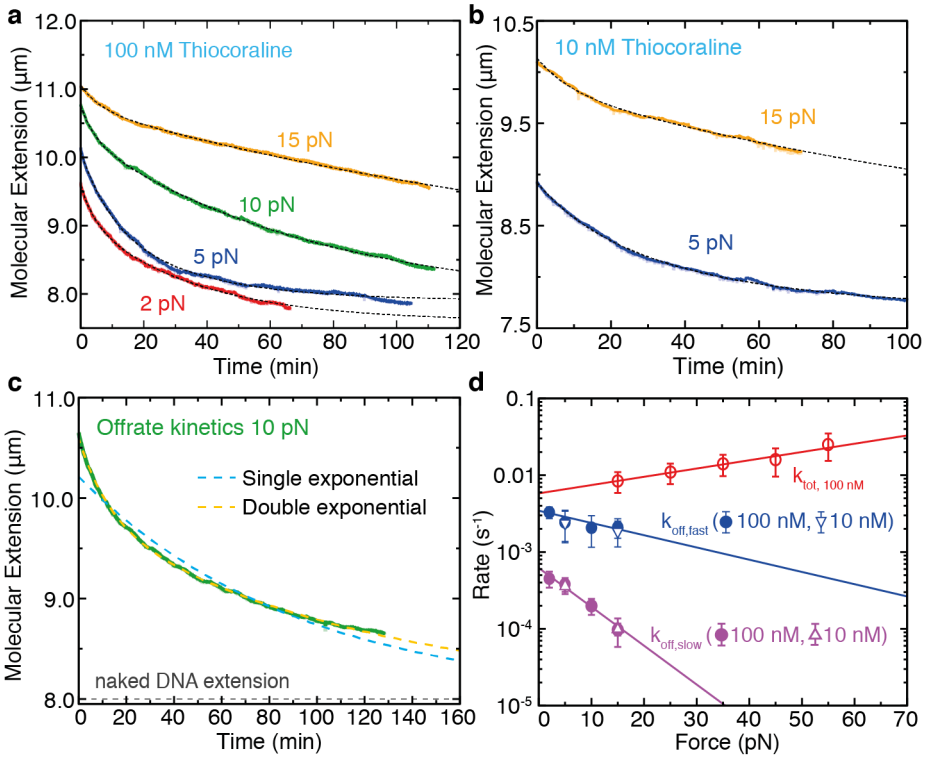

Figure 32: Off-rate kinetics of Thiocoraline. (a,b) DNA extension as a function of time in wash-off experiments starting at a [Thiocoraline] of 100 nM (a) and 10 nM (b). Note the initially shorter molecular extension in the latter. The unbinding reactions are well-described by a double-exponential function (dotted gray). (**c**) Molecular extension as a function of time (green) in a wash-off experiment (initial [Thiocoraline]=100 nM, F=10 pN). A double exponential function (yellow) fits well the data, whereas a single-exponential function (blue) does not. A reference of the extension of a naked dsDNA molecule at 10 pN is indicated in gray. **(d)** Force-dependence of the macroscopic kinetic rates: fast off-rate (blue), slow off-rate (purple). In red the total kinetic rate from force-jump experiments is plotted as a reference. A fit to Eq. 22 (solid lines) yields the macroscopic zero-force rates ($k_{off,slow}(0) = 6.2(7) \cdot 10^{-4}$ s^{-1}, $k_{off,fast}(0) = 3.4(4) \cdot 10^{-3}$ s^{-1}) and transition-state distances ($x^{\dagger}_{off,slow} = -0.48(5)$ nm, $x^{\dagger}_{off,fast} = -0.15(4)$ nm) for each process. Values are reported as mean±SD, N⩾3.

ferent forces even for an initially high concentration of Thiocoraline (100 nM).

8.2 THE KINETIC OFF-RATE (k_{OFF}) REVEALS AN INTERMEDIATE

We investigated the off-rate kinetics of Thiocoraline at different forces in the DNA stretching setup. We first followed the unbinding reaction in DNA molecules equilibrated at 100 nM Thiocoraline, where the initial binding density is large (Figure 32**a**). The off-rate kinetics are remarkably slow at all forces (in the order of hours), showing a strong force-dependence. A single-exponential function fails to fit the experimental data (Figure 32**c**), but experiments are well-described by a double-exponential relaxation whose amplitude A is dominated

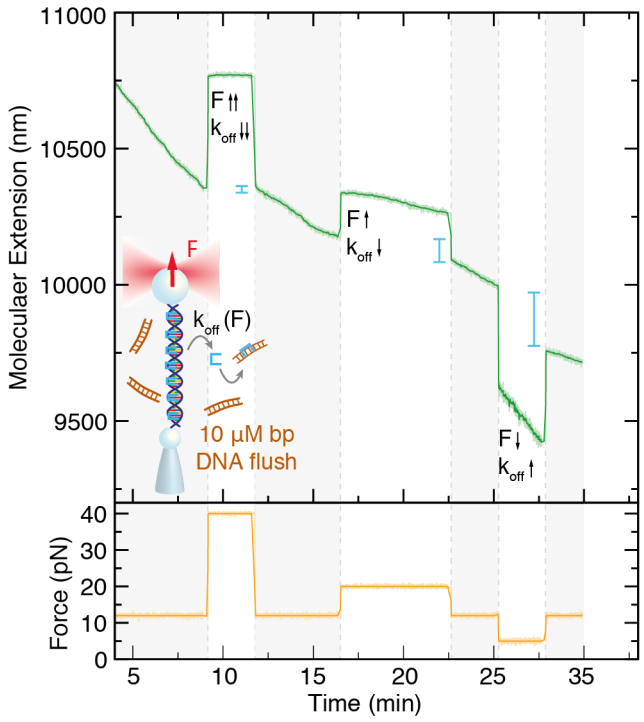

Figure 33: Off-rate kinetics. Wash-off experiment in which different forces (yellow) are sequentially tested within the same kinetic trace. The molecular extension (green) is initially followed at a constant force of 12 pN. Increasing the force from 12 pN (shadowed in gray) to a higher force value (40 pN, 25 pN), causes a decrease on the slope and hence in k_{off}, whereas decreasing the force to 5 pN causes the opposite effect. We also indicated the decrease in molecular extension (blue bars), when the 12 pN force is recovered after a few minutes at 40 pN, 25 pN and 5 pN, showing that ligands unbind faster as the force is decreased. The DNA molecule was initially equilibrated at 12 pN and 100 nM Thiocoraline.

by the slower off-rate process ($A_{slow} > 75\%$, for all forces). The force-dependence of the two off-rates determined in this way are shown in Figure 32**d**.

To verify that the observed off-rates ($k_{off,\,fast}$ and $k_{off,\,slow}$) are independent of the initial binding fraction, we repeated measurements with DNA molecules equilibrated at a much lower Thiocoraline concentration (10 nM, Figure 32**b**), finding rates that are compatible with the ones obtained at 100 nM (Figure 32**d**). The strong force-dependence of the off-rate kinetics can also be observed by changing the applied force in the course of a wash-off experiment (Figure 33). In such experiment, the slope of the molecular extension is proportional to the number of bis-intercalators that unbind from DNA per unit time and hence it is indicative of how the off-rate (k_{off}) changes with force. When force is increased, the unbinding rate slows down (as it can be deduced from an almost flat slope), whereas a force decrease accelerates unbinding (i.e. shows a steeper slope).

What is the origin of the two timescales observed in the wash-off experiments? The presence of two off-rates, together with the previous

observation of an on-rate that is not proportional to ligand concentration, can be rationalized by considering the existence of an intermediate state in the binding reaction. A minimal model in which the binding of Thiocoraline proceeds through a mono-intercalated intermediate state and that considers three microscopic kinetic rates (α_{on}, α'_{on} and α_{off}) is sufficient to reconcile all observations.

We propose a model in which the binding reaction has an on-pathway intermediate state in which only one of the two intercalating moieties is bound to DNA (Figure 34**a**). The model considers two different rates for binding: one rate corresponding to binding one chromophore when the ligand is unbound (α_{on}) that leads to the mono-intercalated intermediate state. This rate is expected to be concentration-dependent. Then there is a second binding rate (α'_{on}) that corresponds to intercalate the second moiety (once Thiocoraline is in the mono-intercalated intermediate state) in order to reach the fully bis-intercalated state. This kinetic rate is expected to be independent of ligand concentration. Finally, we consider that there is a third kinetic rate (α_{off}) that corresponds to the rate at which any of the intercalating moieties de-intercalates. For simplicity reasons, we considered a minimal model in which this unbinding rate is the same both to remove an intercalating moiety from the fully bound bis-intercalated state, or the mono-intercalated intermediate state.The model can be analytically solved (section 8.3), and the values of $\alpha'_{on}(F)$ and $\alpha_{off}(F)$ directly determined from the macroscopic rates $k_{off,\,fast}(F)$ and $k_{off,\,slow}(F)$ measured in wash-off experiments (Figure 34**b**). The microscopic kinetic rates depend exponentially with the applied force F and can be fit to the expression

$$\alpha_{on'/off/on}(F) = \alpha_{on'/off/on}(0)\exp(Fx^{\dagger}_{on'/off/on}/k_BT), \tag{22}$$

where k_B is the Boltzmann constant and T the temperature. Using the microscopic rates determined in this way, the master equations can be solved to reproduce three types of experiments with different initial conditions: (i) wash-off experiments, (ii) force-jump experiments, and (iii) intercalation kinetics starting from naked DNA. Remarkably enough, the amplitudes (A_{slow} and A_{fast}) observed in the wash-off experiments, are in excellent agreement with the amplitudes predicted by the model if the intermediate state is assumed to have an extension 0.4 times that of a bis-intercalated state (Figure 35**a**), supporting the hypothesis that the intermediate state is mono-intercalated (with an elongation ~50% that of the bis-intercalated state which is 0.68 nm). Similarly, the fractional elongation $\chi(t)$ (0 for naked DNA, and 1 for fully bis-intercalated DNA) can be calculated from the molecular extension $x(t)$ of kinetic experiments (see section 8.3) and compared to the prediction of the model. The model describes well the equilibrium extensions and kinetics of the force-jump experiments (Figure 35**b**), and of constant-force experiments that start with a naked DNA

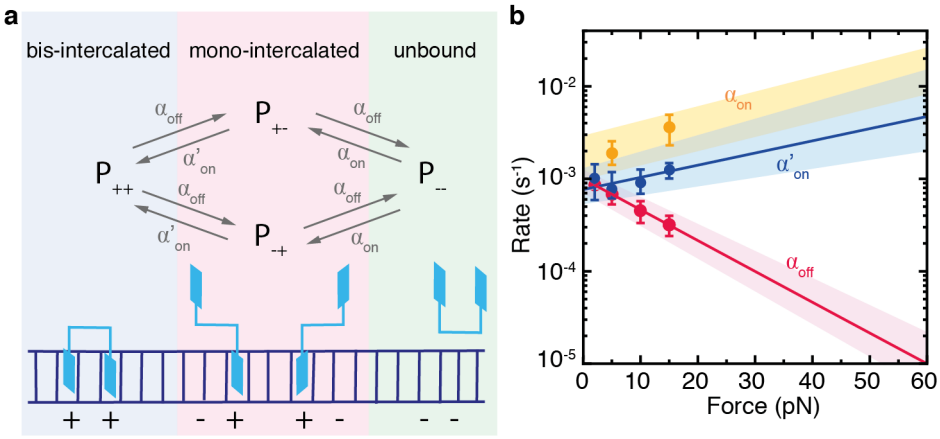

Figure 34: A kinetic three-state model with microscopic rates α_{on}, α'_{on} and α_{off}. (a) Scheme of the kinetic model representing the different accessible states: unbound ligand (green), fully bis-intercalated (blue), and mono-intercalated intermediate (red). The binding of the first intercalating moiety is bimolecular (with a rate α_{on} proportional to ligand concentration), whereas the transition from the mono-intercalated intermediate to a fully bis-intercalated state is unimolecular (with rate α'_{on}). On the other hand, the removal of any of the two intercalative moieties is driven by a single off-rate α_{off}. **(b)** Microscopic rates α'_{on}, α_{on} and α_{off} that best describe the experiments. The values of α'_{on} (blue points) and α_{off} (red points) are determined from the macroscopic rates measured in wash-off experiments (Figure 30**f**). A fit to Eq. 22 (red and blue lines) is used to extrapolate the high force values and their associated uncertainty range (shaded area). The rates α_{on} are determined from intercalation experiments with naked DNA (Figure 35**c** and section 8.3).

molecule (Figure 35**c**). Overall, this three-state model, captures the essential features of the experiments providing a mechanistic explanation for the binding kinetics of Thiocoraline.

The observed increase of α'_{on} with force is characterized by a distance $x^{\dagger}_{on'} = 0.12(5)$ nm, whereas α_{off} steeply decreases with force ($x^{\dagger}_{off} = -0.29(2)$ nm). An off-rate that decreases with force is in striking contrast with previous studies of the mono-intercalator Actinomycin [174] and a threading intercalator [176], were no intermediate was observed. Whereas force facilitates both the binding and unbinding of these ligands due to base pair destabilization [174, 176], this is not the case for Thiocoraline. Our data shows that DNA stretching kinetically stabilizes the bound state, suggesting that bis-intercalators behave like a molecular "wedge" (i.e. applying force clamps the ligand, whereas compression pushes the intercalator out). According to our model, the characteristic elongation of each intercalating moiety can be inferred from the transition-state distances ($x^{\dagger}_{on'}$ and $x^{\dagger}_{off}$). For the transition between the intermediate state and the fully bound state, this yields a value ($\Delta x_{eq} = x^{\dagger}_{on'} - x^{\dagger}_{off} = 0.41(8)$ nm). This is in good agreement with the expected distance change for this transition ~ 0.4 nm (equal to the additional 60% of elongation to reach a full bis-intercalated state). From the value of α_{on} we can also infer the rate of association at zero force ($k_a = \frac{\alpha_{on}(0)}{[Thio]} = 1.7(3) \cdot 10^4\ M^{-1}s^{-1}$). This rate

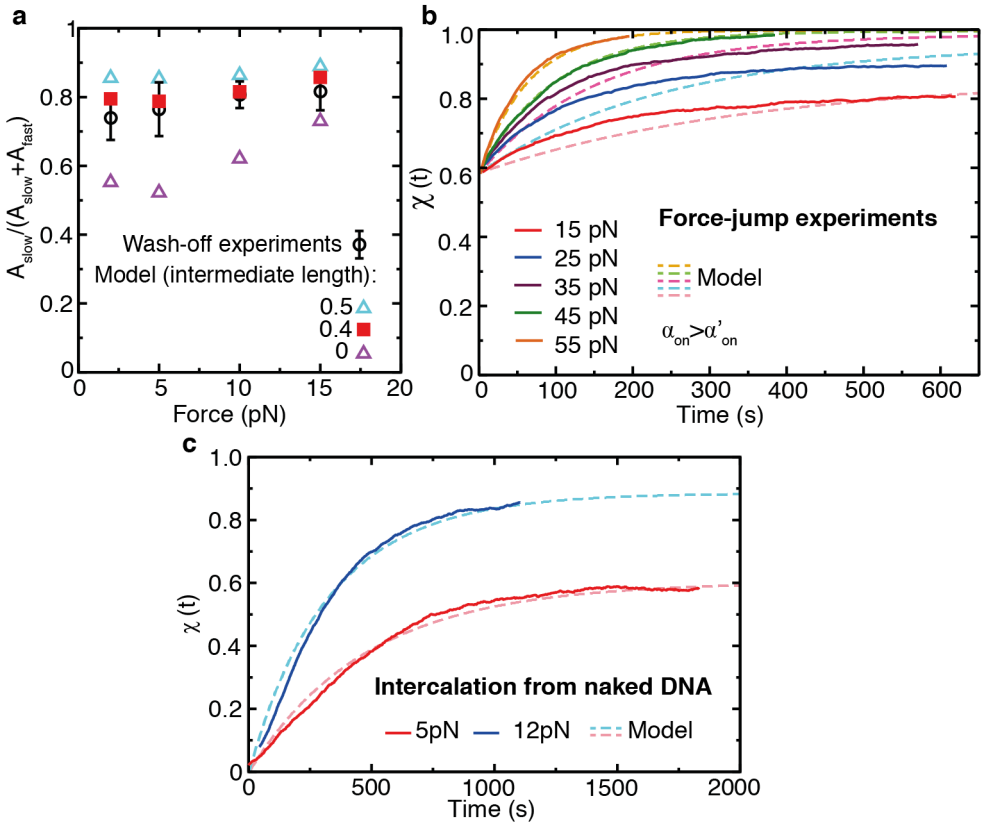

Figure 35: Agreement between experiments and model. (a) Ratio of amplitudes of the fast and slow macroscopic off-rates measured in the wash-off experiments (black) compared with the model prediction (red). **(b)** Fractional elongation of the force-jump experiments from Figure 28**a** (solid lines) and comparison to the kinetics predicted by the model (dashed lines). **(c)** Fractional elongation of intercalation experiments (100 nM Thiocoraline) starting from a naked DNA molecule at 5 and 15 pN (solid lines) and comparison to the kinetics predicted by the model (dashed lines).

is slow if compared to mono-intercalators or charged bis-intercalators (YOYO-1) [191, 192], but is in good agreement with other structurally similar non-charged bis-intercalators [82, 193].

8.3 SOLUTION OF THE THREE-STATE KINETIC MODEL

The kinetics of the model (Figure 34**a**) can be described by means of the following master equations:

$$\frac{dP_{++}}{dt} = \alpha'_{on}(P_{+-} + P_{-+}) - 2\alpha_{off}P_{++} \tag{23a}$$

$$\frac{dP_{+-}}{dt} = P_{--}\alpha_{on} + P_{++}\alpha_{off} - (\alpha_{off} + \alpha'_{on})P_{+-} \tag{23b}$$

$$\frac{dP_{-+}}{dt} = P_{--}\alpha_{on} + P_{++}\alpha_{off} - (\alpha_{off} + \alpha'_{on})P_{-+} \tag{23c}$$

$$\frac{dP_{--}}{dt} = \alpha_{off}(P_{+-} + P_{-+}) - 2\alpha_{onï¿œ}P_{--} \tag{23d}$$

where P_{++}, $P_{+-/-+}$ and P_{--} correspond respectively to the probabilities of a site to be: in the fully bound state (bis-intercalated), in the intermediate state (mono-intercalated), or non-occupied (naked DNA). The occupancy of the intermediate state can be grouped by defining $P_1 = P_{+-} + P_{-+}$:

$$\frac{dP_{++}}{dt} = \alpha'_{on} P_1 - 2\alpha_{off} P_{++} \tag{24a}$$

$$\frac{dP_1}{dt} = 2P_{--}\alpha_{on} + 2P_{++}\alpha_{off} - (\alpha_{off} + \alpha'_{on})P_1 \tag{24b}$$

$$\frac{dP_{--}}{dt} = \alpha_{off} P_1 - 2\alpha_{on} P_{--} \tag{24c}$$

8.3.1 *Determining the microscopic rates α'_{on} and α_{off}*

For wash-off experiments it can be assumed that α_{on}=0 (as no free-ligand is present in the buffer), and the set of equations can be analytically solved by finding the eigenvalues and eigenvectors of the reduced system:

$$\begin{pmatrix} P_{++} \\ P_1 \end{pmatrix} = \begin{pmatrix} -2\alpha_{off} & \alpha'_{on} \\ 2\alpha_{off} & -(\alpha_{off} + \alpha'_{on}) \end{pmatrix} \begin{pmatrix} P_{++} \\ P_1 \end{pmatrix} \tag{25}$$

The eigenvalues of this system are determined as:

$$\lambda_{+-} = \frac{-(3\alpha_{off} + \alpha'_{on}) \pm \alpha_{off}\sqrt{1 + 6\alpha_{off}/\alpha'_{on} + (\alpha_{off}/\alpha'_{on})^2}}{2} \tag{26}$$

The system therefore shows two characteristic timescales (a faster one, λ_- and a slower one λ_+) that correspond to the two off-rates observed in the wash-off experiments ($\lambda_= k_{off,fast}$, $\lambda_+ = k_{off,slow}$), finding that:

$$\alpha_{off} = \sqrt{\frac{k_{off,fast} k_{off,slow}}{2}} \tag{27a}$$

$$\alpha'_{on} = k_{off,fast} + k_{off,slow} - 3\sqrt{\frac{k_{off,fast} k_{off,slow}}{2}} \tag{27b}$$

From the experimental values $k_{off,fast}$ and $k_{off,slow}$ at different forces (Figure 32**d**) we can determine the force-dependence of the intrinsic rates $\alpha'_{on}(F)$ and $\alpha_{off}(F)$ (Figure 34**b**). Using these rates, the master equations can be solved for any set of initial conditions ($P_{++}(0)$, $P_1(0)$ and $P_{--}(0)$), finding solutions that will be of the form:

$$P_{++/1/--}(t) = A_{fast} e^{-k_{off,fast} t} + A_{slow} e^{-k_{off,slow} t} + A_\infty \tag{28}$$

where A_{fast} and A_{slow} are the amplitudes of the fast and slow relaxation modes.

8.3.2 *Comparing model predictions with experimental relaxations.*

In our experiments, we have access to the change in molecular extension as bis-intercalators bind or unbind. From the model depicted in Figure 34**a**, the change in molecular extension $x(t)$ can be related to the probability of the bis-intercalated and intermediate states as:

$$x(t) \propto P_{++}(t) + fP_1(t) = A_{off,fast}e^{-k_{off,fast}t} + A_{off,slow}e^{-k_{off,slow}t} \quad (29)$$

where we have assumed that the intermediate state contributes to a fractional extension, f, of that corresponding to a fully bound bis-intercalator (as we will later see $f \sim 0.4$ maximizes the agreement between model and experiments). Consequently, it is useful to define the normalized fractional elongation of the molecule $\chi(t)$, that goes from 0 (naked DNA) to 1 (maximal elongation due to full bis-intercalation):

$$\chi(t) = P_{++}(t) + 0.4P_1(t) \quad (30)$$

For the wash-off experiments, the kinetics of the system (Eq. 24) can be analytically solved using the $\alpha'_{on}(F)$ and $\alpha_{off}(F)$ values shown in Figure 32**d**. By assuming, that the molecule is initially in the fully bisintercalated state ($P_{++}(0) = 1$, $P_1(0) = P_{--}(0) = 0$), we can determine the expected amplitude of the slow and fast decay rates ($A_{off,fast}$ and $A_{off,slow}$). The ratio between these amplitudes ($A_{off,slow}/(A_{off,slow} + A_{off,fast})$), can be compared to the experimental values obtained, finding compatible results (Figure 35**a**). The agreement between theory and experiment is maximized for an intermediate state that has an extension 0.4 times that of a fully bound bis-intercalator. This value is close to half that of a fully bound bis-intercalator, supporting the hypothesis that the intermediate state is mono-intercalated.

CHAPTER 9

RESULTS: SINGLE-MOLECULE FOOTPRINTING EXPERIMENTS OF PREFERRED BINDING SITES

To determine the preferred binding sites of Thiocoraline, we developed a single-molecule footprinting assay based on DNA unzipping. As we have previously shown in the wash-off experiments, DNA unzipping allows to detect the binding of intercalators due to the increased stability of the DNA duplex at the positions were ligands bind (section 8.1.1). In this section we will show that the position of these binding events can be localized with near base-pair accuracy, and this information used to determine the sequence-selectivity of the ligand. In fact, unzipping-based methods have proven useful to locate binding of nucleosomes and restriction enzymes [194, 195]. However, to the best of our knowledge this is the first time that an unzipping-based method is used to directly determine the binding site of small ligands to DNA.

9.1 LOCALIZING BINDING EVENTS IN DNA UNZIPPING

We repeatedly unzipped a 480-bp DNA hairpin (synthesis details in Appendix D) in the presence of 5 nM Thiocoraline (Figure 36**a**). In the absence of ligand the unzipping and rezipping curves fully overlap (Figure 37**a**, top). However, in the presence of ligand, the unzipping curves show a series of force peaks indicative of individual binding events (Figure 36**a**, marked with *). A statistical analysis of the FDCs allows us to determine the position of these binding events [182, 196].

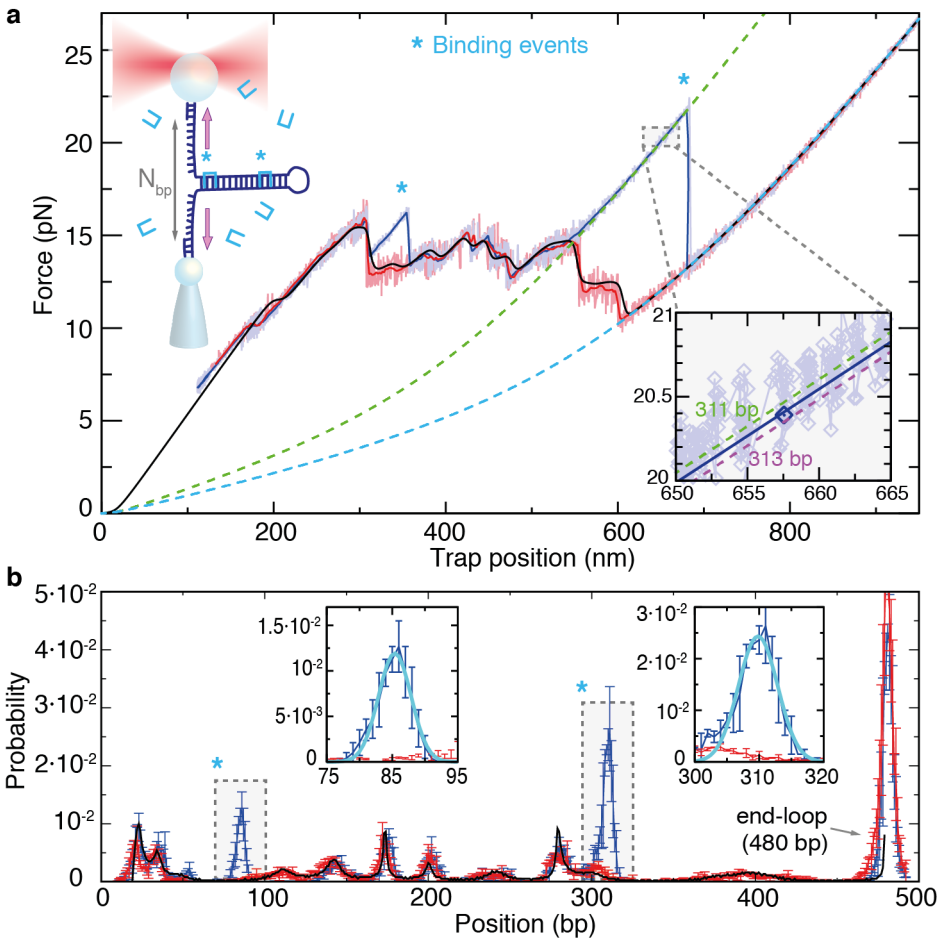

Figure 36: Single-molecule footprinting using DNA unzipping. (a) DNA unzipping (blue) and rezipping (red) of a 480-bp DNA hairpin in the presence of 5 nM Thiocoraline. Pulling speed is 70 nm/s and data is recorded at 1 kHz (light colors) and filtered to 10 Hz (dark colors). The binding of a ligand brings an extra stabilization energy, causing a force peak in the unzipping curve (*). A theoretical prediction of the equilibrium FDC is shown in black and used for alignment. A freely-jointed chain (FJC) curve corresponding to a fully unzipped hairpin (n= 480 bp, cyan dashed) and one that passes close to a binding event (n=311 bp, green dashed) are plotted as a reference. (Inset) For each datapoint of the unzipping/rezipping curves we found the most probable number of open base pairs n, by finding the theoretical curve that passes closest to it. For the highlighted datapoint (blue), the distance is minimized for n=312 bp (blue line). **(b)** Histogram of n values of the unzipping (blue), rezipping (red) and theoretical (black) curves from panel (a). The histograms match each other except for the positions in which there is a binding event (highlighted in gray, *). (Insets) Zoom of a binding peak and fit to a Gaussian function (cyan).

First, the experimental FDCs are aligned to a theoretical prediction based on the nearest-neighbor model (Figure 36**a**, black) [62]. The elastic parameters (l_k, S) used to model the ssDNA are obtained by fitting the relaxation curve of the fully unfolded state (n =480 bp) to the extensible freely-jointed chain (FJC), and are in excellent agreement with previous results for the same ionic condition (chapter A) [197]. A small calibration correction (below 5%) on the experimentally measured distance and force is allowed to fully overlap the theoretical and experimental FDCs between the initial rip (n=27 bp) and the fully unzipped state (n=480 bp).

For every data point (x_{exp}, f_{exp}) of a FDC, we then determine the most probable apparent number of open base pairs (n) by finding the theoretical curve that passes closest to that point at the force f_{exp} (Figure 36**a**, Inset):

$$\left|x_{exp} - x_{th}\left(n, f_{exp}\right)\right| = \min_{m}\left(\left|x_{exp} - x_{th}\left(m, f_{exp}\right)\right|\right). \tag{31}$$

The theoretical extension ($x_{th}\left(n, f_{exp}\right)$) is determined from the different contributions to the total distance:

$$x_{th}\left(n, f_{exp}\right) = x_h\left(f_{exp}\right) + x_b\left(f_{exp}\right) + x_{ssDNA}\left(n, f_{exp}\right) \tag{32}$$

where x_b is the displacement of the bead from the center of the optical trap, x_h the extension of the 29 bp dsDNA handles, and x_{ssDNA} the extension of released ssDNA during unzipping. The term $x_{ssDNA}\left(n, f_{exp}\right)$ is the only term that depends on the number n of unzipped base pairs, and is modeled with a FJC of $2n$ bases with the aforementioned elastic parameters. The other terms do not depend on n and are modeled as: $x_b = f_{exp}/k_{trap}$ (where k_{trap} is the stiffness of the optical trap), and x_h using a WLC (Eq. 56) with the generally accepted parameters for dsDNA [182]. In this way, we assigned a number of open base pairs n to each experimental data point.

The histogram of n for an unzipping curve and the following rezipping curve can then be plotted (Figure 36**b**). FDCs are reversible at this pulling speed, and consequently the unzipping and rezipping histograms (Figure 36**b**, red and blue respectively) overlap except for those positions in which there is a binding event (Figure 36**b** and Figure 37**b**, blue peaks). By fitting these extra peaks to Gaussian distributions, the position of each binding event is determined from the mean value of the Gaussian fit (Figure 36**b**, Insets). A correct matching between the other regions of the histogram is checked and compared to the histogram of the theoretical FDC (Figure 36**b**, black line) to ensure a correct alignment.

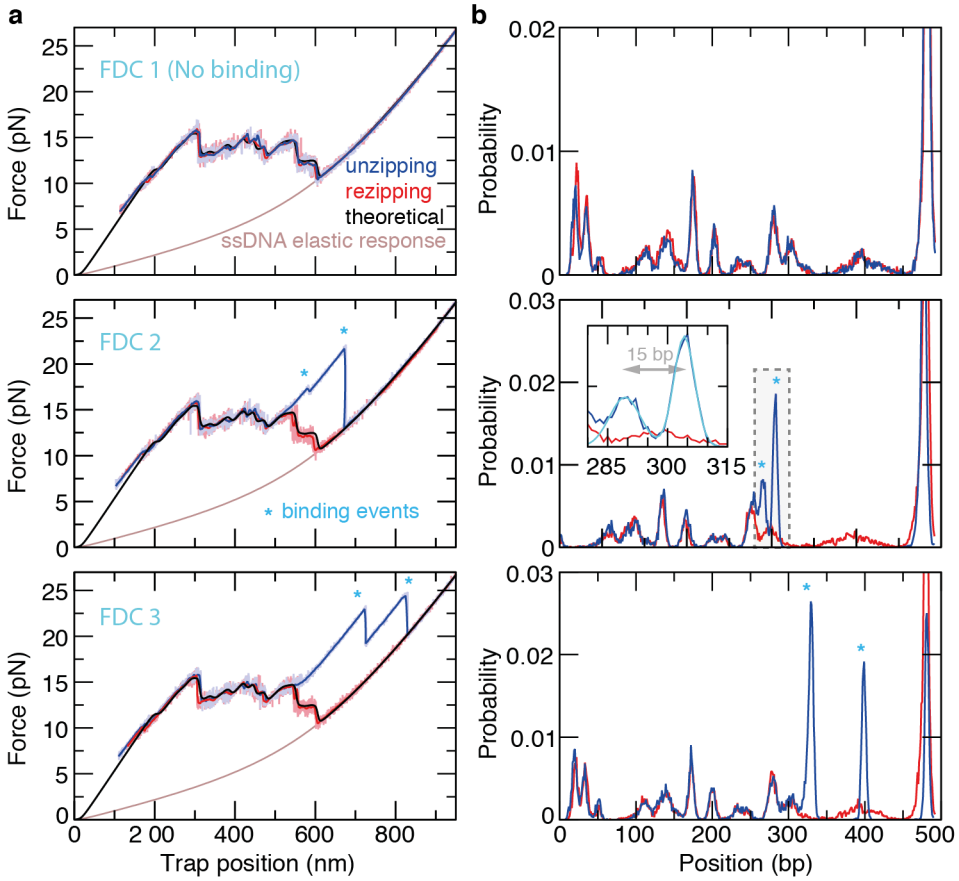

Figure 37: Typical unzipping curves in footprinting experiments. (**a**) Set of three FDCs showing none (top) and 2 (middle and bottom) binding events. Pulling speed is 70 nm/s. At this speed the unzipping (blue) and rezipping (red) curves are fully reversible in the absence of binding events (top). FDC's are aligned to a theoretical curve (black), and a FJC curve of a fully unfolded molecule (480 bp) is plotted as a reference (brown). If an individual Thiocoraline molecule binds to the hairpin, its binding site can be identified from the position of the observed force peak (indicated with an *). Data is obtained at 1 kHz (light blue and red), and a 10 Hz average shown in dark colours. (**b**) Histogram of the number of open base pairs (n) of the FDCs shown in panel (a). As expected, the histograms of the unzipping (blue) and rezipping (red) curve fully overlap for pulling cycles without binding events (top). Binding events are easily identified in the unzipping histogram as peaks that do not have a counterpart in the rezipping curve (middle and bottom, *). These peaks are well described by a Gaussian distribution (cyan, middle inset), and the binding positions determined from its mean. Note for instance, that two binding events closely spaced in the same FDC (middle, inset) can be easily resolved with this technique.

9.2 CORRELATING BINDING EVENTS WITH SEQUENCE PREFERENCES

As the sequence of the DNA hairpin is known (Appendix D), we can correlate the binding positions n determined in this way, with the DNA sequence. The normalized probability of binding for each sequence B_i is determined as:

$$p_{norm}(B_i) = \frac{N_i/S_i}{\sum_i N_i/S_i} \tag{33}$$

where N_i is the number of binding events observed at the sequence i, and S_i is the number of times that the sequence is found in the DNA hairpin (i.e. degeneracy).

In this way, we determined the probability of binding for each different dinucleotide pair (Figure 38**a**, black). Remarkably, we find that Thiocoraline bis-intercalates all 10 possible dinucleotides, but shows a clear preference to clamp CpG steps. Moreover, by analyzing the neighbouring bases we also find that flanking A-T base pairs are favoured (i.e. ACGT, TCGA, ACGA) (Figure 38**a**, Inset). The sensitivity of the technique can be increased by only selecting events that show rupture forces greater than a given threshold (e.g. 5 pN) above the local unzipping force. As expected, this analysis shows a higher peak at the specific binding sites (Figure 38**a**, red), confirming the selectivity of Thiocoraline for clamping CpG steps. An estimation of the measurement error can be found by considering alternative distances between the position of the binding peaks and the clamping positions (Figure 38**b**). By considering different relative distances between the binding peak and the clamping site (i.e. different reading frames), we observe that the sequence-selectivity is only observed for the correct "reading frame" 0. Alternative reading frames, show a distribution compatible with a random binding, and only a low peak is observed for the neighbouring reading frames ± 1. Overall, the fact that a sequence correlation is only clearly observed for the reading frame 0 supports the fact that binding sites are located with ~ one base-pair resolution.

9.3 COMPARISON TO PREVIOUS MEASUREMENTS AND TECHNIQUES

So far, the existence of specific binding sites for Thiocoraline was rather controversial. Although the peptide showed a general preference for G-C rich regions, protected sites were not observed in classical DNA footprinting experiments, suggesting that it had little sequence selectivity [185]. On the other hand, fluorescent melting experiments with short oligonucleotides showed a clear preference for G-C rich oligos and particularly those containing CpG steps [186]. Our direct measurement of binding positions from unzipping ex-

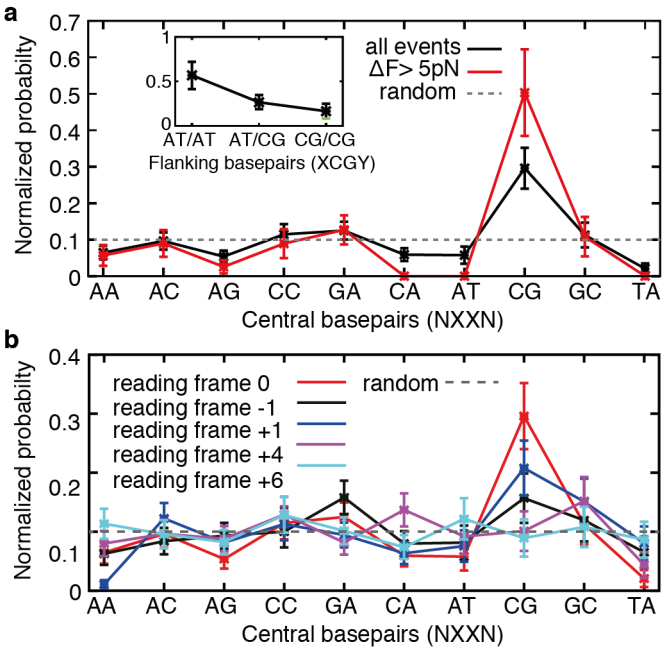

Figure 38: Sequence-selectivity of Thiocoraline clamping positions. (a) Normalized probability of clamping sites XX: including all binding events (black), or by selecting events with rupture forces 5 pN higher than the average unzipping force (red). The dashed line indicates the distribution expected for a ligand that uniformously binds all sites. (Inset) Normalized probability of observing A-T bp or C-G bp flanking the preferred binding site XCGY. (N=142 binding events, 104 unzipping curves and ~1 binding event per cycle). **(b)** Normalized probability of clamping sites XX if different reading frames are considered. Binding at the tetranucleotide position were blockage is observed (reading frame 0, red) is compared to alternative distances from the peak position (black, blue, purple, cyan). Sequence-selectivity is only observed for the reading frame 0, whereas other reading frames show probabilities compatible with a random distribution (dashed gray). Only for the reading frames ±1 (black and blue) a lower peak at CG is also observed, indicative of the error associated with the method.

periments, demonstrates that Thiocoraline has a specific affinity for CpG steps (particularly those flanked by A-T base pairs), although it can also clamp a large number of alternative dinucleotide steps, in agreement with our stretching results that show binding saturation at high enough concentrations (Figure 25**a**). Moreover, this binding pattern seems remarkably similar to that exhibited by the quinoxaline bis-intercalator Echinomycin, that shares several structural similarities with Thiocoraline that might confer specificity for these motifs [76, 81, 198].

Finally, our footprinting assay gives direct access to individual binding events of ligands due to the higher thermodynamic stability that they induce upon binding DNA. Consequently, specific and non-specific binding sites are observed in a single experiment, without relying on indirect measurements such as differential access of cleavage agents, shifts in gel-mobility (difficult for small ligands), and/or radiolabeling [179, 180, 199]. Similarly, DNA templates can be much longer than

typical gel footprinting assays, and a statistically-significant amount of data can be collected in a few hours. We therefore anticipate that this method may be widely applicable to determine the sequence-specificity of a large number of small ligands that are difficult to characterize with traditional assays.

CHAPTER 10

DISCUSSION OF THE RESULTS

Thiocoraline, an anticancer peptide that bis-intercalates DNA, exhibits an unusual slow kinetics of binding. We have exploited this feature to clarify the mechanical properties of bis-intercalated DNA in the low and high-force stretching regimes [87, 171, 178, 188]. Contrary to previous studies in which equilibrium FECs were used, we have worked with reversible metastable FECs in which the fraction of ligand is kept constant. We found that the bending rigidity of DNA remains unchanged, which we interpret in terms of a compensation between the electrostatic effect of unwinding the DNA double helix (larger interphosphate distance) and the straightening induced by base-stacking interactions with the intercalated chromophores.

We also showed that Thiocoraline has a high binding affinity to DNA finding compatible values from equilibrium and kinetic measurements ($K_d = 77 \pm 13$ nM) and reaching the nearest-neighbour exclusion limit of n=4 at submicromolar concentration [200]. Remarkably enough, the affinity to DNA is dominated by a very slow macroscopic off-rate ($k_{off}(0) = 6.2(7) \cdot 10^{-4}\ s^{-1}$) that steeply decreases with force. This allows us to kinetically control drug unbinding by changing the applied force, obtaining extremely long-lived complexes at even moderate stretching forces ($F < 20$ pN). From a structural point of view, this indicates that a pulling force further accommodates the bound ligand, stabilizing the bis-intercalated tetranucleotide motif. This might be common to other cytotoxic DNA bis-intercalators (e.g. Echinomycin, Triostin), and we hypothesize it is very likely related to the high toxicity of this family of anticancer drugs, that show ac-

tivity at concentrations much below their K_d values (e.g. Thiocoraline $IC_{50} \sim 200$ pM [185]). To accurately determine these unbinding rates, we have used a combination of stretching and unzipping experiments that allowed us to ensure that subnanomolar traces of drug did not remain in the wash-off experiments. This might be particularly relevant in future studies of other highly hydrophobic and low solubility ligands [182]. Our experiments also reveal that the binding pathway proceeds through a long-lived mono-intercalated intermediate (lifetime ~ 10 min) which elongates DNA by 0.28 nm (which is 0.4 times the elongation of a fully bis-intercalated ligand, 0.68 nm) contributing to the long lifetime of the peptide (hours). The stretching-unzipping methodology allows us to extract the force-dependence of the microscopic rates ($\alpha_{on}(F)$, $\alpha'_{on}(F)$ and $\alpha_{off}(F)$) from the characteristic timescales of kinetic experiments performed with varying conditions.

To determine the sequence-selectivity of the ligand, we developed a novel DNA unzipping assay to perform single molecule DNA footprinting experiments. Using this assay we could directly locate single binding events finding that Thiocoraline shows a higher preference to clamp CpG steps, particularly those flanked by A-T base pairs. We foresee that this assay might be of extreme utility to determine the binding site of small ligands as it gives direct access to thermodynamically stabilized sites without requirement of cleavage enzymes and radiolabelling techniques, allowing to rapidly recover the binding sites in DNA fragments tenfold longer than typical footprinting templates [179, 180, 199].

So far, Thiocoraline has been observed to interfere with primer extension by DNA polymerase α, showing a very prolonged inhibition of DNA replication even after 48 h removal of the drug [201], in contrast with the typical topoisomerase inhibition mechanism of most DNA intercalators (e.g. doxorubicin). However conclusive studies are lacking, being unclear how its selectivity for cancer cells arises [198, 201]. In view of these results, we suggest that the increased binding stability of the ligand at moderate stretching forces, might create long-lived kinetic roadblocks that stall the progression of replication and transcription enzymes at picomolar drug concentrations. Moreover, we have shown that these roadblocks might also happen through the stabilizaton of non-native kinetic structures that are particularly present in DNA-melted regions during transcription and replication. More importantly, the determined sequence-specificity suggests that Thiocoraline might accumulate at specific genomic locations rich in CpG steps. Notably, CpG sites are sparsely distributed in human genomes, accumulating in many regulatory regions and having much lower frequencies elsewhere [202]. We hypothesize that this inhomogenous distribution of CpG islands might be related to the selectivity of Thiocoraline for carcinogenic cell lines [201].

CHAPTER 11

CONCLUSION

In the work explained in this chapter, we show for the first time that Thiocoraline bis-intercalates DNA via a long-lived intermediate (lifetime ~10 min) that confers unusual properties to the intercalated DNA. To experimentally demonstrate the existence of the intermediate, we designed novel wash-off experiments where the peptide was removed in the fluidics chamber by flowing scavenger DNA. This allowed us to characterize the unbinding kinetics of Thiocoraline at different forces ensuring that traces of ligand did not remain in the chamber. De-intercalation shows a two-step (fast and slow) relaxation process that steeply decreases with force, where most of the relaxation signal (~70%) comes from a long-lived intermediate where only one of the moieties binds DNA. We exploit the large lifetime of the intercalative DNA-Thiocoraline complex to determine the elastic properties (persistence length and stretch modulus) of the intercalated DNA from pulling experiments. Till now this kind of measurements (previously carried out with other mono and bis-intercalative molecules) was not feasible as the DNA intercalative state changed with force along the pulling cycle, due to the much faster binding kinetics, leading to inaccurate conclusions. The right combination of elastic measurements and kinetic experiments allows us to characterize the equilibrium dissociation constant and microscopic rates of the three-state intercalative reaction, providing a model and a powerful methodology applicable to other cases. Finally, by doing unzipping-based footprinting assays we have discovered a specific binding mode of Thiocoraline, that preferentially clamps CpG di-nucleotides along

the DNA sequence. This unzipping assay turns out to an accurate, reliable and fast footprinting method capable of identifying binding motifs of small ligands to DNA with one-base pair resolution. Overall the results of this work, strongly suggest that this specific binding mode together with the slow unbinding kinetics, plays a key role on the high toxicity of this and other cytotoxic bis-intercalators (e.g. Echinomycin, Triostin) at sub-nanomolar concentrations much below their equilibrium K_d values (~80 nM for Thiocoraline). Ultimately our results strongly suggest that the interspersed CpG regions along the DNA crucially enhance the selectivity of Thiocoraline for carcinogenic cell lines.

CHAPTER 12

FUTURE PERSPECTIVES

The combination of DNA stretching and unzipping assays explained in this part of the thesis opens exciting perspectives to determine the binding pathways of biomedically relevant drugs. In particular, the applications of the single-molecule footprinting technique presented in chapter 9 holds exciting prospects. To develop a single-molecule high-throughput footprinting platform, we have adapted the experimental design explained in chapter 9 to a magnetic tweezers instrument, and present it here as a future perspective. As a proof-of-principle of the technique we show its application with the bis-intercalator Echinomycin, which has well established preferred binding sites determined from classical gel footprinting assays [81, 76, 198].

As explained in section 1.2.1.3, using magnetic tweezers it is possible to simultaneously track >100 beads, greatly simplifying the parallelization of measurements. This allows to achieve high statistics in very short timescales (~ hours). However, the experimental design and footprinting protocol must be adapted to the characteristics of magnetic tweezers setups. Particularly, magnetic tweezers are a natural force-clamp, and unzipping experiments must be performed by controlling the force applied to the molecule instead of the trap position. In a typical DNA unzipping experiment with magnetic tweezers, the force is ramped from a low force value (where the hairpin is folded) up to a high force value (where the hairpin is unfolded). As DNA unzipping happens in a narrow range of forces (~13-16 pN), the hairpin unfolds in a cooperative fashion in this force range (i.e.

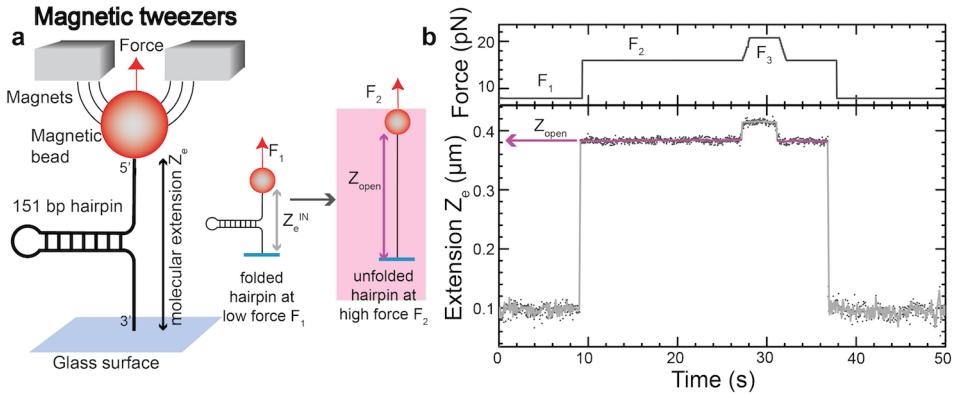

Figure 39: Unfolding a DNA hairpin with magnetic tweezers. (a) Experimental setup. **(b)** A DNA hairpin can be cooperatively unfolded by changing the force from a low force value (F_1), to a high force value (F_2). The large change in molecular extension (Z_e) corresponds to the unfolding of the 151 bp of the hairpin. The extension of the fully unfolded hairpin at the force F_2 is highlighted in purple (Z_{open}). Further increasing the force to F_3, causes a small change in extension due to the elastic stretching of the ssDNA.

a sudden increase in molecular extension is observed). If the protocol is reversed and the force decreased to its initial value, the hairpin folds back to its native structure. Such pulling cycles typically show a large hysteresis due to the cooperative unfolding (refolding) of large DNA segments when force is increased above (decreased below) the critical force of the hairpin (~13-16 pN), and are not well suited to apply the methodology explained in chapter 7. However, a footprinting protocol can be developed by performing force jumps.

Figure 39**a** shows the basic experimental setup, in which a DNA hairpin is tethered between a magnetic bead and the glass surface. At a low force (F_1) the hairpin is in the folded configuration and has a short molecular extension (Figure 39**b**). If the force is increased to a high force value (F_2), the hairpin unfolds in a single step reaching a much longer extension indicated as Z_{open} (Figure 39**b**, purple). If the force is then increased to a value F_3, a small change in extension is observed due to the compliance of the ssDNA. The hairpin can be folded back to its native state by decreasing the force to its initial value (F_1), recovering the initial molecular extension.

In order to perform footprinting experiments with magnetic tweezers we developed a protocol in which these three force values (F_{1-3}) are alternated (Figure 40**a,b**). First the force is jumped from the low force ($F_1 \sim 6$ pN, hairpin folded) up to a force high enough to unfold the full hairpin ($F_2 \sim 17$ pN). This force (F_2) is chosen to be high enough to unfold the hairpin in a single step in the absence of ligands (as previously shown). However, if a ligand (e.g. Echinomycin) is bound to the hairpin, a temporally blocking event that delays the unfolding of the hairpin is observed at the position were the ligand is bound (three blocking events -red, blue and green- are observed in Figure 40**b**). The position of these blocking events can be deter-

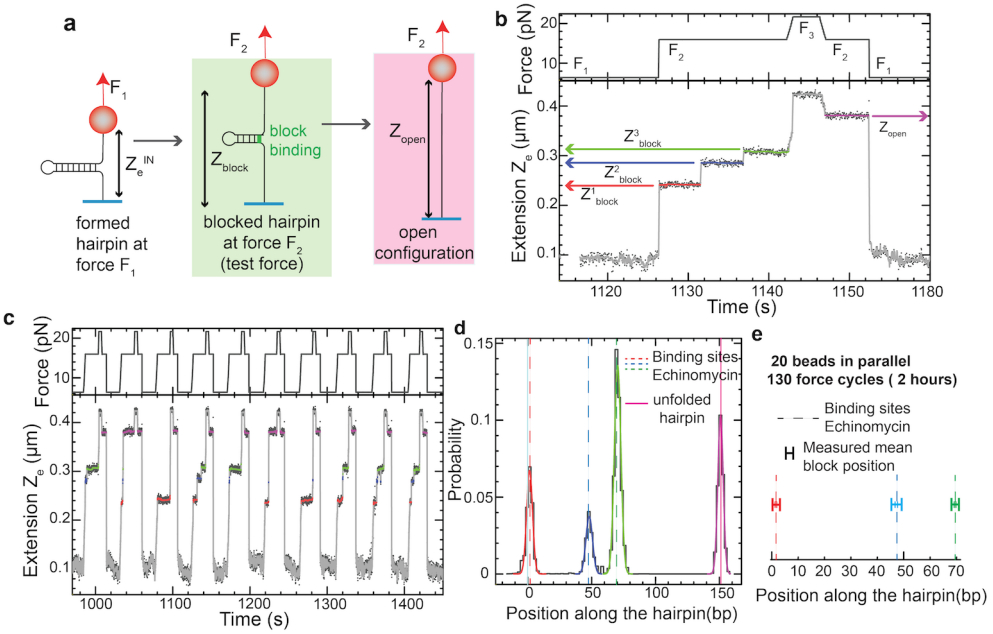

Figure 40: High-throughput footprinting using magnetic tweezers. (a) Experimental setup. **(b)** Protocol for DNA footprinting using MT. Three force steps (F_{1-3}) are used to repeatedly unfold the DNA hairpin and detect the position of binding events (Z^{1-3}_{block}, red, blue, green) in relation to the fully unfolded state (Z_{open}, purple) **(c)** The protocol is repeated several times in order to obtain high statistics of the preferred binding sites. **(d)** Histogram of blocking events observed for the bis-intercalator Echinomycin. By using the method explained below, molecular extensions (in nm) are transformed to positions (in bp), and the preferred binding sites are determined by fitting the peaks to Gaussian functions. **(e)** Mean position of the binding peaks ($\langle z_1 \rangle = 2.0$ bp, $\langle z_2 \rangle = 47.4$bp, $\langle z_3 \rangle = 69.2$ bp, with $\sigma \sim 2.7 - 3.2$ bp) compared to the expected binding sites for Echinomycin at 2, 47 and 69 open bp (dashed lines).

mined from the measured extension Z_e if the blocking events are long enough ($\sim$ 1 s.). After $\sim$20 s, the force is increased to the higher force value ($F_3 \sim 25$ pN) to remove any remaining ligands bound to the template and reach the fully unfolded configuration. The force is decreased to its initial value in two steps: first it is maintained for a few seconds at the testing force (F_2) and then decreased down to the low force value (F_1) to refold the DNA hairpin. The intermediate step at F_2 is performed to obtain the reference extension (Z_{open}) of the fully unfolded hairpin at the same force at which blocking events are observed (Figure 40**b**, purple). This measurement is used to determine the location of the binding events relative to the end of the hairpin (in nanometers). By repeating this footprinting protocol several times (Figure 40**c**), statistics of the preferred binding sites of the ligand can be collected. The blocking events determined in this way can then be plotted in an histogram (Figure 40**d**), and the peaks fitted to a Gaussian function. The position of the binding peaks (red, blue, green) are aligned with the reference peak corresponding to the unfolded hairpin (Figure 40**d**, purple). From the mean value of the Gaussian

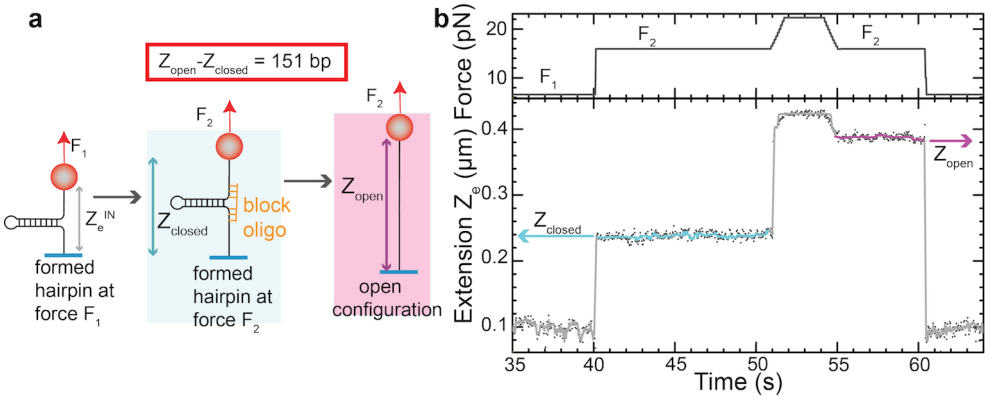

Figure 41: Conversion of measured extensions (nm) into basepairs positions. (a) Experimental setup to measure the extension of the fully folded (Z_{closed}) and unfolded (Z_{open}) state at the testing force F_2 using an oligonucleotide complementary to the ssDNA handles of the hairpin. **(b)** In the presence of the blocking oligonucleotide, the unfolding of the hairpin at the force F_2 is delayed. In this way, the reference extension of the folded hairpin (Z_{closed}, cyan) and the unfolded hairpin (Z_{open}, purple) at the force F_2 can be simultaneously measured. As the hairpin contains 151 bp, the number of open base pairs of binding peaks can be directly determined.

fits, the preferred binding sites of the ligand are determined with one basepair resolution (Figure 40**e**).

The conversion of the measured distances from nanometers into basepairs is achieved by performing measurements with a blocking oligonucleotide that delays the opening of the hairpin structure at F_2 (Figure 41**a**). This provides a reference measurement of the extension of the unfolded hairpin with respect to the folded hairpin (($Z_{open} - Z_{closed}$, 151 bp)) at the same forceF_2 at which ligand binding is tested (Figure 41**b**). These references can be used to set a conversion factor in bp units at the force F_2, from which the position of the binding peaks is then determined. These measurements with the blocking oligonucleotide can be performed simultaneously to the footprinting experiment, providing a precise measurement of the extension of the fully folded (Z_{closed}) and fully unfolded (Z_{open}) hairpin for each tracked bead.

Proof-of-principle experiments have been performed with a hairpin of 151 bp with a sequence that contains binding sites for several small ligands with varying binding modes (groove binders, intercalators), recognition sequences (AT-rich, CG-rich), and affinities (pM-µM). The sequence of the hairpin is engineered to have a flat free energy landscape in its initial 120 bp (Figure 42). This is done to ensure that the observed blocking events are exclusively due to the extra stabilization energy brought by the bound ligands, and not by the DNA sequence itself (e.g. kinetic barriers due to CG-rich regions). Indeed, the development of DNA hairpins that contain all the possible tetranucleotide combinations but that at the same time have a flat free-energy landscape is an important development to make single-molecule footprinting with magnetic tweezers fully competitive with

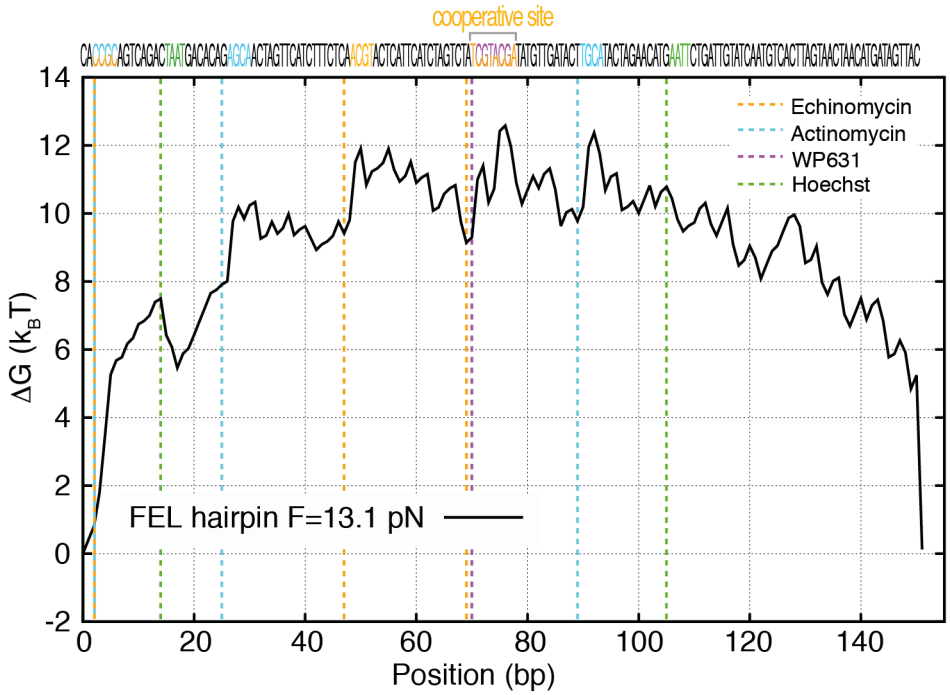

Figure 42: Sequence and free-energy landscape of the hairpin. Free-energy landscape of the 151 bp hairpin used for the footprinting experiments at its critical force (F ~ 13.1 pN). The free-energy is mostly flat and does not show peaks larger than 2-3 k_BT (in the order of thermal fluctuations) along the central region of the hairpin. The location of potential binding sites of a series of small ligands is indicated in colored dashed lines. The sequence of the DNA hairpin is indicated above, and potential binding sites are indicated using the same color code. Overlapping sites are indicated with two colors. Note as well the existence of two juxtaposed sites for the bis-intercalator Echinomycin (position 69), where the ligand is expected to show cooperative binding [83].

other techniques. Consequently, although the parallelization of measurements is a strong advantage of magnetic tweezers, the requirement of a flat free-energy landscape, might make optical tweezers a preferred method to footprint specific DNA sequences and ligands. A throughout comparison between both single-molecule footprinting techniques (e.g. studying the ligands depicted in Figure 42) would be very helpful to make single-molecule footprinting useful to the chemical biology community. As a long term perspective, the possibility to develop assays in which epigenetic modifications (e.g methylations), or reconstituted nucleosomes are incorporated, would provide information currently inaccessible with other techniques.

Part III

FORCE-SPECTROSCOPY OF PEPTIDE-DNA AGGREGATION AND COMPACTION

"The union of opposites, in so far as they are really complementary, always results in the most perfect harmony; and the seemingly incongruous is often the most natural."

Stefan Zweig (Ungeduld des Herzens)

CHAPTER 13

INTRODUCTION

13.1 MOTIVATION

The formation of proteinaceous aggregates has been observed since the pioneering experiments on the physical chemistry of proteins by Fischer and Hofmeister in the first years of the XXth century. In fact, aggregation is tightly coupled to protein folding and stability, and it was not until 1911 that Chick and Martin first distinguished the process of aggregation from that of denaturation [203], setting the bases for the forthcoming experiments that unveiled protein secondary and tertiary structures and its thermal stability. Indeed, whether proteins were formed by long polypeptide chains or by colloidal aggregates of small peptides remained controversial for almost 30 years, until the advent of Svedberg's ultracentrifuge in 1924 which allowed to determine that proteins are formed by highly homogenous populations of large molecules [204].

For a long time, protein aggregation and the formation of inclusion bodies (i.e. large cytoplasmatic accumulations of protein aggregates) remained as a disturbance to by-pass during the overexpression of proteins for *in vitro* studies, and little attention was paid to its relevance. However, in the last 30 years there has been a growing recognition that the prevalence of protein aggregation is much higher than generally thought. This renewed interest has been due to several reasons. On the one hand, the growing concern on diseases related to protein aggregation (e.g. Alzheimer's and Parkinson's diseases, prion diseases) lead to the discovery of the beta-amyloid (Aβ) in 1985 [205]

and boosted studies on the formation and structure of amyloid fibrils and its relation to disease. On the other hand, the major economic impact of protein and compound aggregation in the biotechnological (e.g. large-scale protein purification [206, 207]) and pharmaceutical industries (e.g. nonspecific inhibition of target molecules by small drug aggregates [156], immunogenic response against aggregated drugs [208], drug delivery [209]), has also encouraged research on the principles behind molecular aggregation in order to develop more accurate prediction and profiling tools for early stages of drug development [210]. Finally, and in a more imaginative fold, their use as building blocks for biotechnological applications such as electrochemical devices has also been recently proposed [211, 212, 213].

A main challenge in investigating aggregates lies in their high heterogeneity: aggregates may differ in size (from a few nm to mm), morphology (amorphous or fibrillar), type of intermolecular bonding (covalent vs. non-covalent), reversibility, and in the case of proteins also structure (native vs. denatured) [115]. This inherent complexity has been addressed during the last decades with the development of new techniques that have proved very useful to characterize the structure and kinetics of formation of these species in a reproducible way [121, 149, 214]. For example, developments such as solid-state NMR (ssNMR) spectroscopy and advances on X-ray crystallography have dramatically improved our knowledge of amyloid structure [215, 216]. Still, a proper understanding of their mechanism of action requires not only knowing how aggregated species are formed but also tackling down how they interact with other molecules present in their biological context (e.g. in the cellular milieu or the extracellular matrix) [119, 129, 217, 157].

For instance fibrils obtained from amyloid deposits in *in vivo* samples are usually found complexed with polyanions (especially polysaccharides and nucleic acids). *In vitro* experiments have shown that indeed, polyanions electrostatically bind to amyloid peptides (e.g. human lysozyme, Aβ-40, α-synuclein) and promote their aggregation by reducing their electrostatic self-repulsion [107, 129, 130, 131, 132, 218, 219, 220]. Even if it is still unclear how these complexes might relate to disease, such ubiquitous interaction has triggered discussion on its connection to neurodegenerative disorders, and also on the hypothetical role of aggregating peptides as scaffolds for polynucleotide assembly in early prebiotic life [129, 221]. Yet, similar questions are also being faced by the pharmacological industry, where colloidal aggregates of organic compounds non-specifically bind and sequester molecular targets through hydrophobic interactions, affecting the results of high-throughput assays [156, 157, 222]. Similarly, latest experimental evidences relate amyloid cytotoxicity to the binding of early unstructured oligomers to the plasma membrane through exposed hydrophobic patches [147]. Overall, a deeper understanding of the

forces that drive self-aggregation and how these compete with binding to other molecules is becoming a topic of increasing interest in a variety of fields (from physical chemistry and biochemistry to pharmacology and nanotechnology).

In this sense, it is important to recall that two main forces controlling these interactions are hydrophobic and electrostatic interactions. The former trigger self-assembly of molecules with exposed hydrophobic residues due to the "hydrophobic effect" (see section 4.2); the latter prevent self-association of -locally or globally- highly charged proteins, and favor interactions with macromolecules which exhibit a high compensatory charge. It is therefore a trade-off between these two main forces that largely drives both self-association and molecular interactions of aggregating species [117, 129, 223].

Unfortunately, the complex and nonspecific nature of these interactions makes difficult to quantitatively determine key parameters such as the binding affinities and kinetic steps during the interaction process. Indeed, a full characterization of the interaction between hydrophobic peptides and polyanions in aqueous buffer is especially challenging due to the competition between peptide-peptide self-aggregating interactions and peptide-substrate electrostatic binding forces, as well as due to the transient and heterogenous nature of the formed complexes [132, 223].

Consequently, one of the main focus of this thesis has been precisely to obtain quantitative information of such interactions by using single-molecule force spectroscopy techniques. This part of this thesis (Part II) describes how optical tweezers can be used to dissect the kinetic steps involved in the interaction between an hydrophobic anticancer agent with a high tendency to aggregate -Kahalalide F-, and a highly charged polyanion such as DNA. The selection of this 13-residue peptide, which contains a single positive charge, allowed us to begin to differentiate between electrostatic and hydrophobic effects in the complexation process (see section 13.2 for details on Kahalalide F). As a polyanion, DNA is an excellent template to test such interactions in force-spectroscopy experiments due to its well-established chemistry for optical tweezers manipulation. Finally, and as a future perspective, we also show that similar experiments can be conducted with more complex peptides such as the amyloid $A\beta$-40 peptide, and consequently an introduction on amyloid structure and formation is also presented in section 4.1. From a physical point of view, understanding how this trade-off between electrostatic and hydrophobic forces regulates self-assembly and binding appears as a challenging and exciting goal.

13.2 KAHALALIDE F

13.2.1 *Isolation and structure*

In the last decades the highly competitive environment of the sea has arised as a powerful source of bioactive compounds (e.g. antitumoral, antifungal, antiviral and anti-inflammatory molecules) that can be used to treat a large number of human conditions [224]. These highly cytotoxic compounds are usually found as secondary metabolites in small invertebrates and algae that use them as a first-line of protection against fish predation.

Kahalalide F (KF) is a 13-residue cyclic depsipeptide originally isolated from the Hawaiian mollusk *Elysia rufescens* and its diet *Bryopsis sp.* (Figure 43) [225]. KF and its derivative elipsidipsin (Irvalec®) are the most promising compounds of the family of the kahalalides due to their potent cytotoxic activity against several tumour cell lines [226, 227, 228]. The peptide (Figure 43**b**) contains several non-proteinogenic modifications such as two non-coded amino acids (ornithine -Orn- and dehydroaminobutiric acid -Dhb-), and amino acids with D-stereochemistry (e.g. D-isoleucine) [229]. It has an overall complex chemical structure that is composed by:

- A 6-amino acid cyclic part closed by an ester bond that links the carboxylic acid of a L-Valine (C-terminus of the peptide) and the secondary alcohol of a D-Threonine;
- and a 7-amino acid exocyclic tail that ends with a saturated aliphatic acid at the N-terminus (methylhexanoic -MeHex- tail).

Altogether, KF has a highly hydrophobic structure (contains 10 hydrophobic amino acids and the MeHex tail), and just a single protonable group (L-ornithine) that is positively charged at neutral pH ($pK_a \sim 10.76$). This makes KF a low solubility compound with a high tendency to self-assemble and aggregate in aqueous solution. Still, and as we will later see, the single positive charge of the ornithine residue also allows KF to establish electrostatic interactions with negatively charged substrates such as nucleic acids.

13.2.2 *Mechanism of action and activity*

Compared to other anticancer drugs, KF presents a novel mechanism of action that 20 years after its first isolation is still far from being fully understood. Although new promising strategies based in antibody therapy are in the pipeline [230, 231], most current anticancer agents effect their cytotoxic activity by inducing apoptosis (i.e. programmed cell death) of carcinogenic cells either through activation of cell death receptors on the membrane [232, 233], by interfering with

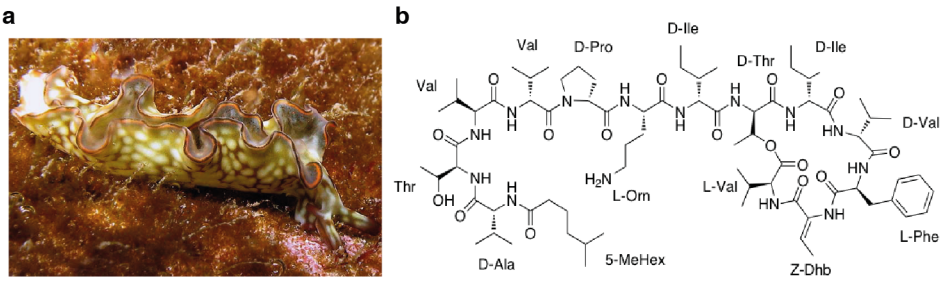

Figure 43: Kahalalide F. (a) The nudibranch *E. rufescens* from which Kahalalide F was isolated. Picture by F. Michenet. **(b)** Kahalalide F structure. The peptide is mainly composed of nonpolar and hydrophobic amino acids, being the ornithine residue (L-Orn) the only protonable group (MW: 1477.87 g/mol).

the cytoskeleton dynamics [234], or by producing DNA damage and interfering with the DNA machinery [112].

KF instead induces cell death through a process that does not show any of the typical features of apoptosis, but whose main fingerprints are dramatic cell swelling, vacuolization, and severe alteration of lysosomal membrane morphology [226, 227, 228] (Figure 44**a**). It also induces mithocondrial damage, plasma membrane rupture and endoplasmatic reticulum vesiculation [227]. In contrast, the nuclear membrane seems to remain unaltered and DNA fragmentation (typical of apoptotic cell death) is not observed in KF-treated cells, although chromatin clumpling into small condensates has been reported [227, 235]. Moreover, the mechanism of action of KF appears to be caspase-independent and does not show other hallmarks of apoptosis such as cell cycle arrest [227]. As previously stated, these results suggest that KF does not induce apoptosis as most anticancer agents [237], but that the peptide induces a necrosis-like cell death whose specific molecular mechanism remains to be clarified. Moreover KF shows a high selectivity for tumour cells over normal non-cancer cell lines [227], indicating that it targets a specific molecular property of carcinogenic cells that is not exploited by current anticancer agents. Although the peptide is safe and showed activity in patients with solid tumours in phase I clinical trials, its development has been recently halted due to the low incidence of these tumours.

The high hydrophobicity of the peptide together with the large effects observed on membranes of KF-treated cells, has encouraged research on finding molecular mechanisms somehow related to the cellular membrane [228]. Some groups have reported that KF interferes with several signalling pathways (ErbB3, EGFR, TGF-α, PI3/AKT) [235], finding that the sensitivity of cancer cells to KF strongly correlates to the expression levels of the membrane receptor tyrosine kinase ErbB3 [238, 239]. This cell surface receptor is involved in the epithelial-mesenchymal transition (i.e. a mechanism by which epithelial cells loose adhesion to the extracellular matrix and are able to migrate

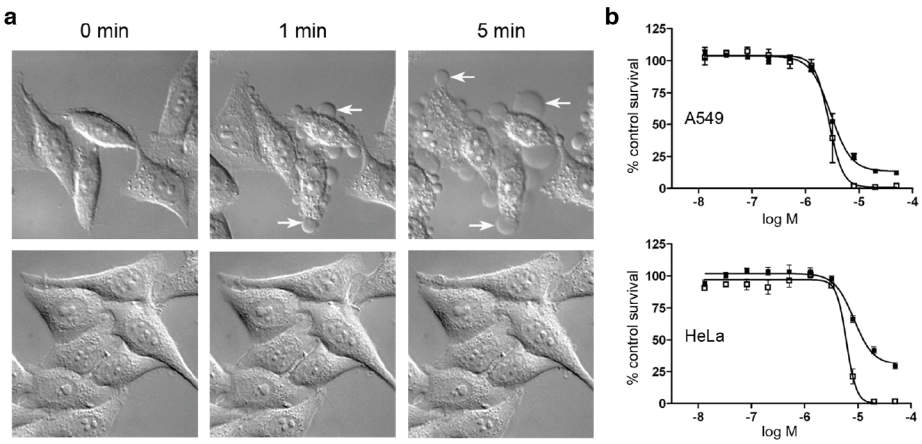

Figure 44: KF cytotoxicity. (a) Three phase-contrast video-microscopy images of HeLa cancer cells treated with 10 μM (top images) and 1 μM (bottom images) Irvalec® at different treatment times. At the cytotoxic concentration cell swelling and blebbing are clearly seen (arrows). **(b)** Dose-response cytotoxic curves for two different cancer cell lines after 30 min (black) and 72 h (white) treatment. Irvalec® shows an all-or-none mechanism and rapidly kills cancer cell above a critical concentration (~ 10 μM). log(M) represents the logarithm of the peptide concentration. Reproduced from [236].

to other organs), essential in the process of tumor progression and metastasis [240, 241, 242]. Although these cell surface receptors are a good indicator of response to KF treatment, it has been recently shown that this sensitivity is an indirect consequence of the overall reorganization that KF induces on the structure of the plasma membrane [243]. Indeed KF causes alterations on the order and fluidity of the plasma membrane immediately after administration (Figure 44**a**). The peptide rapidly inserts and permeabilizes the membrane altering the cellular homeostasis and its conductivity [236]. It is remarkable that these alteration are accompanied by cytotoxic curves with a very step dose-response: whereas at 1 μM KF does not cause any visible effect, at 10 μM cells show the characteristic membrane disruption and blebbing and end up with necrotic cell death after just a few minutes (Figure 44**b**). Such an all-or-none mechanism is characteristic of cooperative processes in which molecules must self-assemble into oligomers or higher order structures (see section 4.1.1.2). Indeed, fluorescence experiments also showed that KF oligomerization is a required step for its cytotoxicity [236]. At cytotoxic concentrations (10 μM) the peptide is found to self-assemble in the plasma membrane disrupting it, whereas at low concentrations (1 μM) it evenly distributes into it without visible effects. The formation of these KF oligomers was mainly observed on the plasma membrane, but also in some locations near the nucleus [236].

Such nonspecific membrane disruption mechanism might seem at odds with the strong selectivity of KF for cancer cells. It is known, though, that carcinogenic cells have characteristic changes in their

phospholipid, glycoprotein and glycolipid composition (i.e. proteins or lipids containing covalently-linked oligosaccharides) compared to normal healthy cells [244, 245, 246, 247]. Even if a clear target has not been identified so far, it might well be that one or several of such modifications are responsible of its high selectivity for carcinogenic cells. As a final remark, although our work does not focus on the mechanism of action of the peptide, to our knowledge it is the first one to investigate in detail its aggregation, which has just very recently seen to be at the core of its cytotoxicity, and particularly on how it interacts with a highly charged polyelectrolyte such as DNA. Indeed, our results on its aggregation kinetics and complexation with nucleic acids suggest that KF aggregates form oligomers that could interact in *in vivo* conditions with polyanionic molecules present on the extracellular matrix (e.g polysaccharides) or certain phospholipids.

CHAPTER 14

RESULTS: INTERACION OF KAHALALIDE F WITH dsDNA

In this chapter we describe the interaction between the hydrophobic peptide KF and dsDNA. We first performed a set of experiments in the optical tweezers set-up in which we show how KF electrostatically binds to dsDNA inducing its condensation. We then show how the hydrophobicity and self-aggregating properties of the peptide affect the formation of the complex, leading to a phenomenology that has no counterpart in previous force-spectroscopy DNA condensation experiments. To rationalize the experimental results we build a simple theoretical framework that takes into account this interplay between electrostatic binding forces and hydrophobic self-aggregating forces, and perform further experiments to validate the hypothesis. The single-molecule results are complemented with AFM images and DLS measurements that give further insight on the complexation process. Finally a simple theoretical model is used to simulate the experimental FECs and determine important parameters of the interaction such as the spectrum of binding affinities, kinetic barriers and lengths of DNA segments sequestered within the complex.

14.1 KF COMPACTS dsDNA

To study how KF binds DNA we stretched a single half λ-DNA (24-kb, synthesis details in Appendix D) in the presence of KF in the optical tweezers set-up. First a DNA molecule was tethered between two beads and its elastic properties measured using the WLC model

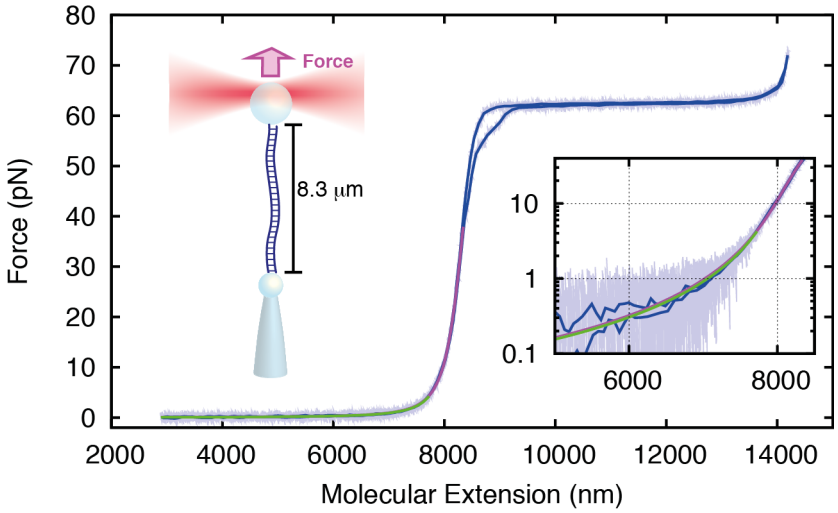

Figure 45: DNA pulling curve of a 24-kb dsDNA molecule at 100 mM NaCl. The FEC (blue) is well described by the inextensible WLC model at low forces (green) and up to a higher force of 40 pN by introducing a stretch modulus (magenta). The force plateau at ~62 pN corresponds to the overstretching transition. Data is recorded at 1 kHz (light blue) and smoothed to 10 Hz bandwidth (dark blue). **(Inset)** Log-scale representation emphasizing the force-region relevant for the WLC fits.

(Figure 45)[248]. The persistence length (l_p) and stretch modulus (S) of each dsDNA molecule were determined before flowing KF finding average values of $l_p = 44.7 \pm 2.0$ nm and $S = 1419 \pm 240$ pN (N=10), in accordance with the generally accepted parameters [249, 250, 251]. A compatible value of $l_p = 43.0 \pm 2.0$ nm was also found using the inextensible WLC model (for details on elastic models and fitting see Appendix A). The stretching curves also showed the characteristic overstretching plateau at a force of 62.5 ± 0.5 pN with an extension ~70% of the contour length. Molecules that showed an abnormally high hysteresis on the overstretching transition (generally attributed to highly nicked DNA molecules) were discarded.

Next the DNA molecule was rinsed with 40 μM KF while it was maintained at an end-to-end distance of 6 μm. In this configuration the tethered molecule can explore bended conformations due to thermal fluctuations, as the force remains below 0.4 pN at this extension. The flow was temporarily stopped after 5, 15 and 30 min in order to record a series of FECs (Figure 46**a**).

After flowing KF for 5 min, DNA maintained at low tension was compacted by the peptide. In order to stretch the compacted molecule, the KF-DNA complex must be unraveled, and therefore a sawtooth pattern with many force rips was observed (Figure 46**a**, purple). This suggests that KF behaves as a DNA condensing agent, inducing kinks and loops on the DNA. The relaxation curves however do not show peaks. In fact, they were similar to those obtained for naked DNA (Figure 46**a**, black) indicating that DNA compaction only took place

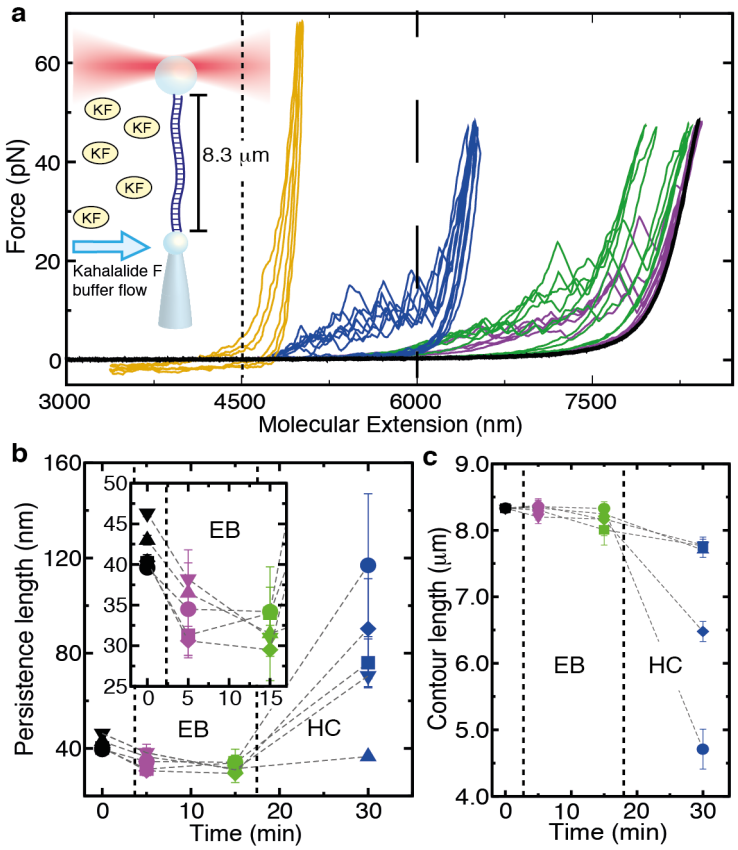

Figure 46: KF binds to dsDNA. (a) DNA pulling curves before (black) and after flowing KF at different waiting times: 5 min (purple), 15 min (green) and 30 min (blue). The molecule is maintained at an extension of 6 μm (vertical dashed line), and the flow is temporarily stopped to perform pulling cycles between a minimum extension of 4.5 μm (dotted line) and a maximum force of 45 pN. The sawtooth pattern observed indicates that KF induces the compaction of DNA. Pulling cycles reaching end-to-end distances lower than 4 μm are shown in yellow. Data is filtered at 10 Hz bandwidth, v=500 nm/s. **(Inset)** Experimental set-up. **(b,c)** Persistence and contour length of five DNA molecules after flowing KF (black corresponds to naked DNA). The changes in the elastic parameters are a signature of the two regimes observed in KF-DNA complex formation: electrostatic binding and hydrophobic collapse.

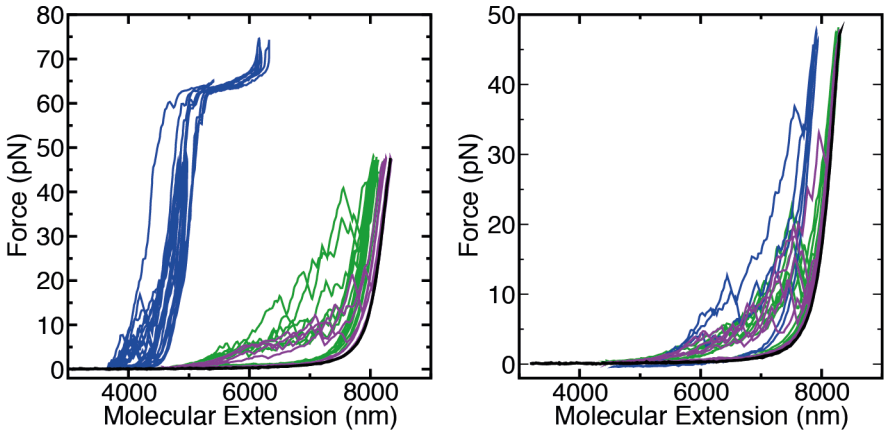

Figure 47: Reproducibility of KF-DNA FECs. Pulling cycles of a 24-kb DNA molecule before (black) and after flowing 40 μM KF at different waiting times: 5 min (purple), 15 min (green), 30 min (blue). Similar trends to those reported in Figure 46**a** are seen. Data is filtered at 10 Hz bandwidth, v=500 nm/s.

after the extension of the molecule was reduced (i.e. compaction lags behind the elastic response of the DNA). Consequently, the relaxation curves were well described by the WLC model, but showed a persistence length 25% lower (Figure 46**b**, Inset, purple) than that of naked dsDNA. This reduction of the persistence length is likely due to the positively charged ornithine residue that decreases the self-repulsion of the DNA phosphate-backbone.

In experiments where KF was flowed for 5 or 15 min, the whole contour length of the DNA could be recovered after pulling up to 40 pN (Figure 46**c**). However, after 15 to 30 min, many interactions could not be disrupted leading to an apparent shorter contour length, that was correlated to an increase of the persistence length. These results suggests that the KF-DNA complex started to collapse into a more stable and stiffer structure after 15 min. Remarkably, a repulsive negative force was detected after 30 min in most of the experiments if end-to-end distances lower than 4 μm were allowed (Figure 46**a**, yellow), suggesting the formation of a thick KF-DNA aggregate of 1-3 μm of length. It was not possible to remove the bound KF by rinsing the molecule with peptide-free buffer for more than 45 min, reflecting the high stability of the final complex. We collected measurements for at least ten different molecules finding a reproducible pattern (Figure 47).

Binding of KF to DNA exhibits two regimes. First, there is a weak and fast regime apparently determined by the electrostatic attraction between the positively charged residues of the KF particles and the negatively charged backbone of DNA. According to our interpretation in this regime DNA binds to hydrophilic spots exposed on the surface of KF particles. Unpeeling of DNA segments requires forces

typically lower than 20 pN. We will refer to this mode of binding as electrostatic binding (EB). This regime is observed in the first 15 min of the experiments shown in Figure 46**a**, and is characterized by a constant contour length, the presence of force rips associated to unpeeling events and a reduced persistence length. The increased flexibility of the filament indicates a charge compensation that reduces self-repulsion of phosphates along the DNA backbone. There is a second stronger binding regime that occurs over longer timescales, that we attribute to the formation of an increasing number of stable hydrophobic contacts within the growing KF-DNA complex. Our interpretation is that in this regime DNA gets buried inside the bulk of the aggregate after being recruited by the hydrophilic spots exposed on its surface creating a stiff filament (as suggested by its increased persistence length). The slower timescales observed in this regime suggest that this mode of binding requires the remodeling and growth of a strongly hydrophobic complex. We will therefore refer to this second regime as hydrophobic collapse (HC). It is characterized by an irreversible decrease in contour length, an increase in persistence length, and the final collapse of the KF-DNA complex that is eventually compressed by pushing the two beads closer than 4 μm. The force required to disrupt such HC structure is above those accessible with our set-up (~ 100 pN), in agreement with previous AFM pulling experiments of single hydrophobically collapsed polymers.[252]

What is the parameter that controls the prevalence of each binding regime? We expect that DNA bending fluctuations determine its binding to KF and the subsequent stabilization of the complex. Therefore, the molecular extension -or distance between beads- should be the parameter controlling the transition between both regimes. To verify this hypothesis we carried out experiments where the DNA was repeatedly pulled in the presence of KF between a maximum force of 40 pN and a minimum extension that progressively decreased from 8 μm to 2 μm in steps of 500 nm per pulling cycle (Figure 48**a**). Such minimum extension controls the degree of compaction reached by the complex. For each cycle we then measured the apparent contour length at the maximum force (40 pN). The results (Figure 48**b**) confirm the presence of the aforementioned regimes, which are separated by a threshold capture distance of 5.5 μm (corresponding to 66% of the contour length of the molecule). In the EB regime (relative extension $\geqslant$ 66%) the apparent contour length of the DNA fiber does not change whereas in the HC regime (relative extension $\leqslant$ 66%) it decreases linearly with the minimum distance between the two beads.

The previous experiments provide evidence that the initial binding of the peptide to DNA is mediated by electrostatic interactions. One would therefore expect the interaction to be highly dependent on the ionic strength of the buffer; as increasing ionic strengths have a higher concentration of counterions that cause a greater screening

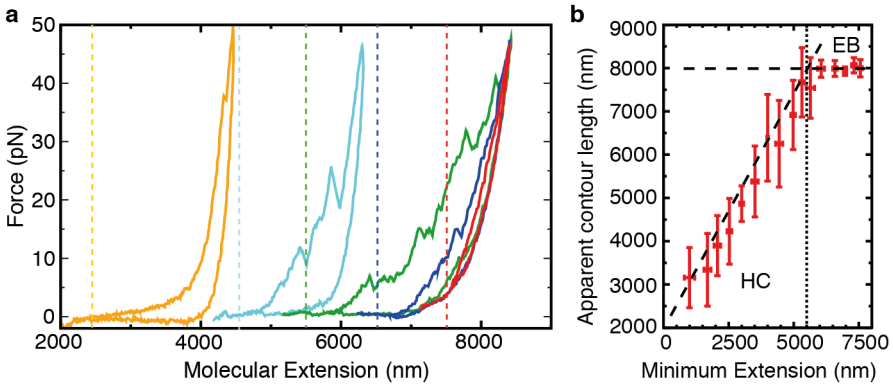

Figure 48: Experiments with decreasing minimum extension. (a) Traces of a typical pulling experiment in which the minimum DNA extension is decreased at every puling cycle in 500 nm steps. Each dotted line represents the minimum extension reached before the shown FEC (color code). Only 5 pulling cycles are shown for clarity. **(b)** Apparent contour length of a DNA molecule repeatedly pulled between a maximum force of 40 pN and a minimum extension that decreases in steps of 500 nm per pulling cycle (mean±SD, N=10).

of electrostatic interactions (i.e. the Debye length is increased). So far we have only investigated the binding of KF to DNA at a single salt concentration (100 mM NaCl) which is near to physiological conditions. Consequently, to confirm our interpretation we performed experiments at varying NaCl concentration (10 mM - 500 mM). The salt titration shows that the initial binding of KF to DNA and the subsequent DNA compaction is highly dependent on the ionic strength of the buffer, in agreement with the proposed mechanism. Specifically, it can be seen that:

- the forces needed to unravel the complex decrease with increasing ionic strength (Figure 49**a-d**), and
- the level of compaction of the DNA (its apparent contour length) decreases with increasing ionic strength (Figure 49**e**);

both facts being indicative of a weaker interaction when the ionic strength of the buffer is increased. Indeed, at the highest ionic strength (500 mM NaCl) we did not observe binding of the peptide to DNA, except for one experiment in which we obtained the results shown in Figure 49**d**.

Finally, to unambiguously establish the relation between the only charged group of the peptide (ornithine) and the binding to DNA, we investigated a KF analog in which the ornithine residue was replaced by a negatively charged glutamic acid. The characteristic sawtooth pattern induced by KF is not observed after flowing the analog, and if FECs are fitted to the WLC model the obtained elastic parameters are compatible with those of naked dsDNA. The fact that no interaction between the KF analog and DNA was observed (Figure 49**f**), confirms

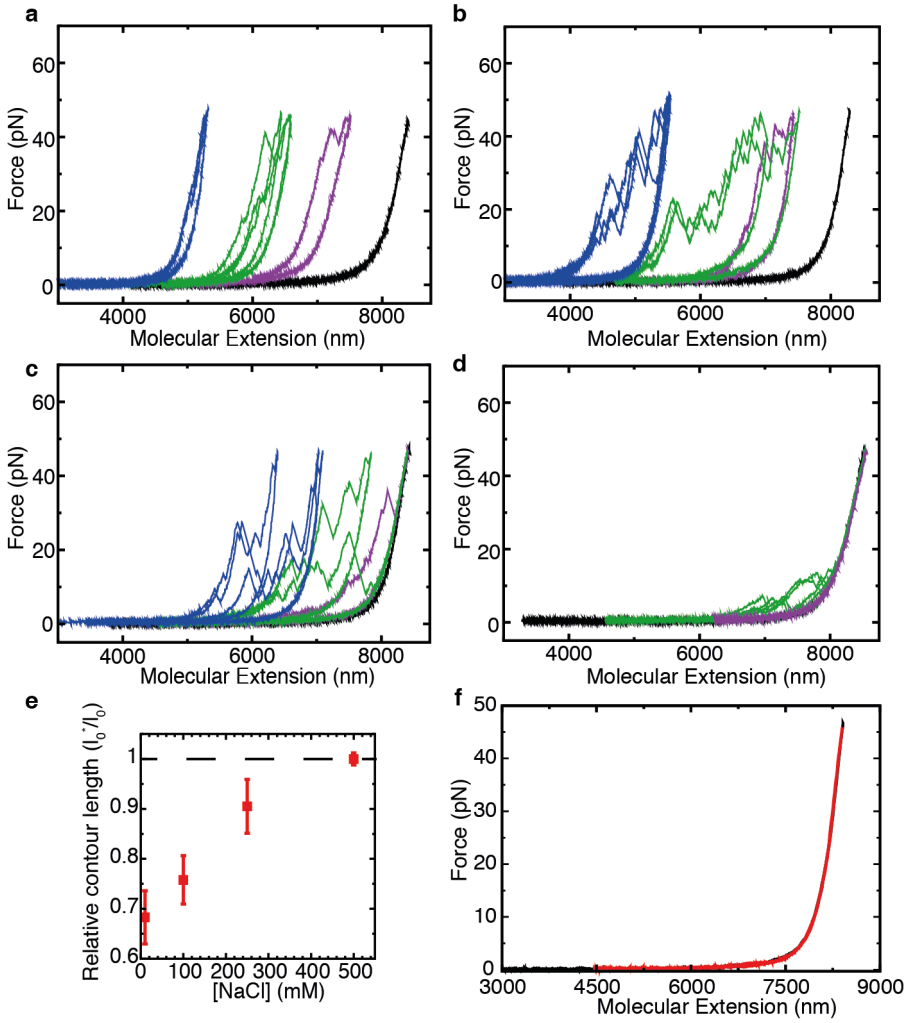

Figure 49: Electrostatics of KF-DNA interaction. Effect of ionic strength: (a) 10 mM NaCl. (b) 100 mM NaCl. (c) 250 mM NaCl. (d) 500 mM NaCl. In panels (a-d) we show representative experiments at each salt condition in which a DNA molecule is pulled before (black) and after flowing 10 μM KF (color). The DNA molecule was repeatedly pulled between a maximum force of 45 pN and a minimum extension of 6 μm (purple), 4 μm (green) and 3 μm (blue). Data is filtered at 50 Hz bandwidth, pulling speed v=500 nm/s. We performed at least 5 experiments at each condition. **(e)** Apparent contour length of the DNA molecule (l_0*) relative to its original extension (l_0=8.3 μm), after being pulled between a maximum force of 45 pN and a minimum extension of 4 μm. The degree of compaction of the DNA molecule increases with decreasing ionic strength (mean±SD, N=5). **(f)** DNA pulling curves before (black) and after (red) flowing a KF analog (40 μM, 100 mM NaCl) in which the ornithine residue has been replaced by a glutamic acid.

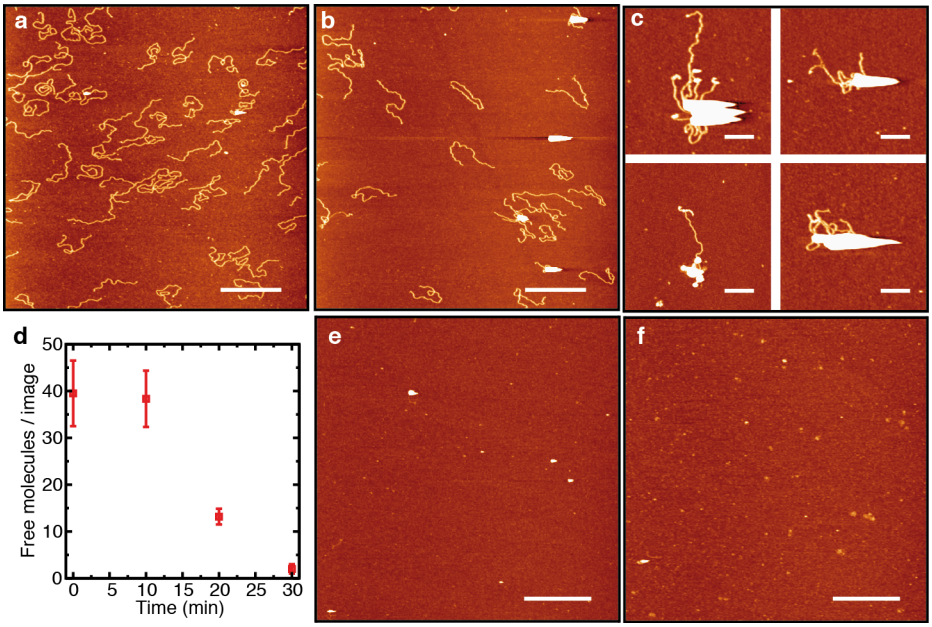

Figure 50: AFM images of KF-DNA complexes. (a-c) AFM images of reactions of 1.65 ng linearized pGEM plasmid (2743-bp) and 100 µM KF obtained at 0, 20 and 30 min incubation times at room temperature, respectively. The number of free DNA molecules decreases with incubation time and large compaction blobs are observed. **(d)** DNA surface density at different incubation times, determined as the average number of free individual molecules per image of 9 μm^2 (mean±SD, N⩾6). **(e, f)** AFM images of KF in the absence of DNA. 100 µM KF immediately after dilution in aqueous buffer (e), and after 30 min incubation at room temperature (f). Small aggregation spots are occasionally observed on the surface at both incubation times. Bar scale is 600 nm. Bar scale is 600 nm (a, b, e, f) and 200 nm (c). Color scale (from dark to bright) is 0-2 nm in all AFM images.

that the positive charge of this residue is essential for electrostatic binding and provides further evidence that electrostatic interactions are key for the initial binding of the peptide to DNA.

14.2 AFM IMAGES OF KF-dsDNA COMPLEXES

The formation of KF-DNA aggregates was directly observed by AFM imaging using a 2743-bp DNA fragment on spermine-treated mica (Figure 50**a-c**). Formation of blobs was observed at the initial time of mixing, and their average size and number increased with time. Notably, after 20 min, a sharp decrease in the number of individual molecules bound to the mica surface was observed (Figure 50**d**). We attribute this to the formation of intermolecular complexes in which several DNA molecules are recruited into a single aggregate. As a consequence, no free DNA was observed after 30 minutes incubation time.

In AFM images of KF samples without added DNA, we did not observe large peptide aggregates bound to the mica surface, and smaller

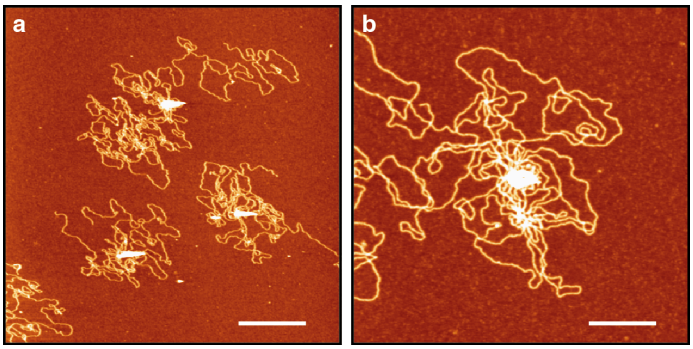

Figure 51: AFM images with a long DNA (48-kb). (a, b) Full λ-DNA incubated with 100 μM KF for 30 min. DNA condensation and formation of blobs are also seen for this larger DNA substrate. Bar scale is 600 nm (a) and 200 nm (b). Color scale (from dark to bright) is 0-2 nm.

ones (i.e. smaller than those observed during KF-DNA complexation) were found to bind in a non uniform manner as large areas appear free of peptide aggregates (Figure 50**e,f**). This non homogeneous adsorption can also contribute to decrease in the density of DNA molecules with time, as large complexes of DNA and KF may be adsorbed in non-imaged regions. Moreover, the capability of DNA to bind to mica may be reduced when complexed with KF. In any case, KF-DNA interactions clearly resulted in aggregation and compaction of DNA molecules supporting the optical tweezers data. These results were further confirmed with the use of longer DNA molecules, in which the formation of compaction spots and intermolecular complexes is clearly observed (λ-DNA, 48-kb, Figure 51).

14.3 DLS MEASUREMENTS OF KF-dsDNA COMPLEXES

The observed aggregation of KF and its complexation with DNA was also characterized with Dynamic Light Scattering (DLS). DLS is a bulk technique widely used to determine the size of sub-micrometrical particles in a solution by measuring their Brownian motion (e.g. proteins, micelles, polymers). This is done by measuring the decorrelation time of the fluctuations of light scattered by the solution when a laser beam is passed through it. Tiny particles diffuse faster through the solution than larger ones and lead to faster fluctuations on the scattered light. Therefore by measuring the autocorrelation time of the light scattered by the sample, it is possible to determine the diffusion coefficient of the particles and hence its hydrodynamic radius (see Appendix C.3 for a detailed explanation).

We first measured the size of KF particles in peptide dilutions and studied how they evolved over time without addition of DNA. The measured autocorrelation function shows that KF aggregates are found in the dilutions immediately after preparation (Figure 52**a**).

Moreover, the increasing autocorrelation time of the measurements (position of the shoulder in Figure 52**a**) indicates that their size also increased with time. From these measurements the hydrodynamic radius of the aggregates can be determined using the Stokes-Einstein relationship (Appendix C.3). The hydrodynamic radius of KF particles one minute after sample preparation was 170±20 nm (Figure 52**b**, red). The hydrodynamic radius increased linearly with time with a growing rate of 3.2±0.6 nm/min, indicative of a rate of aggregation proportional to the surface of the aggregate. We attribute such growth to the hydrophobic interactions between peptides that favor the formation of a solid phase due to the hydrophobic effect. The samples, though, have a high polydispersity (i.e. the aggregates are heterogenous in size). This is inferred from the long tails observed in Figure 52**a**, that causes a discrepancy between the experimental measurements and a fit to a single exponential decay (Figure 52**a**, dotted lines) at long autocorrelation times. This effect becomes more prominent at longer incubation times, when the aggregates grow in size. As the sample shows a distribution of sizes (and not a single size), the autocorrelation function is better described by a distribution of decay rates than by a single exponential decay. Assuming that the distribution of sizes is monomodal, the width of the distribution of sizes can be determined by fitting the autocorrelation data to a function that also considers a distribution of decay rates (i.e. cumulant analysis) [253]:

$$g_2(\tau) = B + \beta e^{-2\Gamma\tau}\left(1 + \frac{\mu_2}{2!}\tau^2 + \mathcal{O}(3)\right). \tag{34}$$

where B is the baseline ($B \sim 0$), β the instrumental response, Γ the mean decay rate, τ the autocorrelation time, and μ_n higher order moments (see Appendix C.3). The error on the size measurements was determined from the second moment (μ_2) of a fit to Eq. 34, which is related to the width of the distribution.

We then performed measurements of KF-DNA mixtures. In these experiments a DNA concentration was added to the KF sample immediately after preparation. Interestingly, the hydrodynamic radius of KF-DNA mixtures remained constant within experimental errors during 60 min after sample preparation (Figure 52**a**, black). This may be explained by the stabilizing effect induced by the added DNA. In fact, the average size of the particles could be stabilized by adding DNA at a latter time-stage, and not immediately after preparation (Figure 52**c**, arrow). DNA as a strongly charged polyelectrolyte interacts with the positive charge of KF and might form an anionic, water-soluble shell around the peptide.

Finally, we also measured the zeta potential of KF particles finding a low positive value (Table 2) -too low to prevent aggregation-, but suggesting that peptide aggregates might expose positively charged residues on the surface. The addition of DNA to the sample shifts the

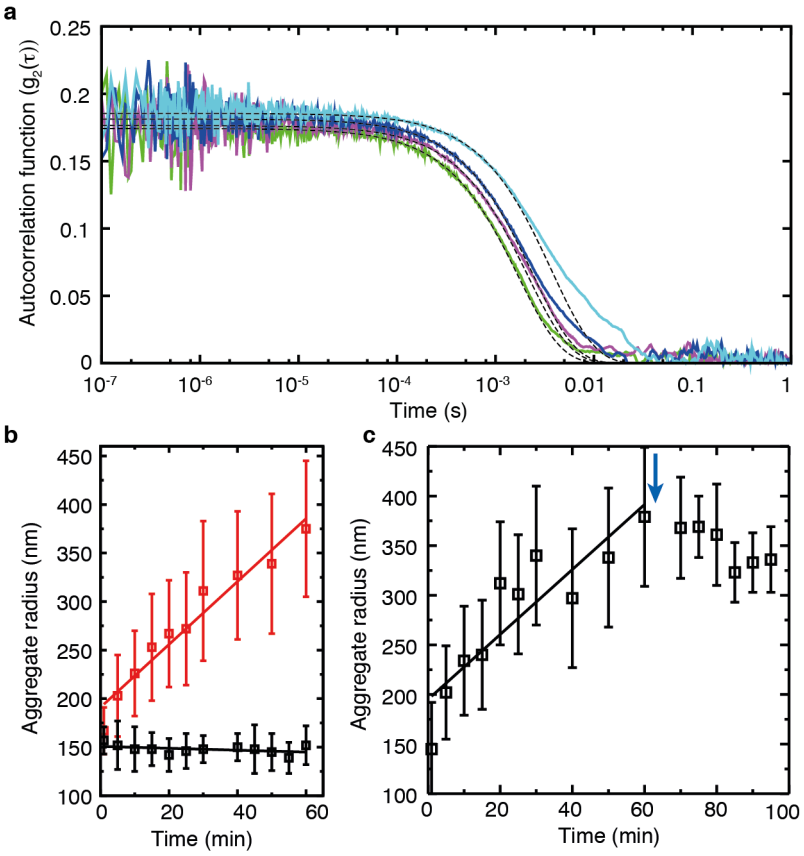

Figure 52: DLS measurements of KF and KF-DNA complexes. (a) Autocorrelation function of a KF dilution measured 5′ (green), 15′ (purple), 25′ (blue) and 45′ (cyan) after sample preparation. The hydrodynamic radius of the particles can be determined from a fit to an exponential decay using Eq. 63 (dotted lines). The increasing autocorrelation time indicates particle growth. **(b) (red)** Hydrodynamic radius of KF particles in the buffer used for optical tweezers experiments (40 μM KF, 25°C, mean±SD, N=9). KF forms nanometer-sized aggregates whose size grows linearly with time. The aggregation rate is obtained from a linear fit (red line). In black, hydrodynamic radius of KF aggregates when DNA is added to the sample immediately after dilution (KF 40 μM, 48-kb λ-DNA 6.25 μg/ml, 25°C, mean±SD, N=3). The size of the KF aggregates remains constant up to 1 hour after dilution, showing that DNA has a stabilizing effect on the size of the aggregates. A linear fit is shown in black. **(c)** The size of the KF aggregates (black) can also be stabilized by adding DNA (arrow) after 60 min (mean±SD, N⩾3). A linear fit is shown to highlight the stabilization of the particle size after adding DNA.

	ZETA POTENTIAL (mV)	ELECTROPHORETIC MOBILITY ($\mu m \cdot cm \cdot V \cdot s^{-1}$)
KF	7.3 ± 0.3	0.6 ± 0.1
DNA	-46.1 ± 5.5	-3.4 ± 0.3
DNA/KF	-30.1 ± 1.2	-2.4 ± 0.1

Table 2: Zeta potential measurements. Zeta pontential of KF (40 μM) and λ-DNA (6.25 μg/ml, 0.2 nM) solutions as well as their mixtures in water at 25°C (mean±SD, N⩾3).

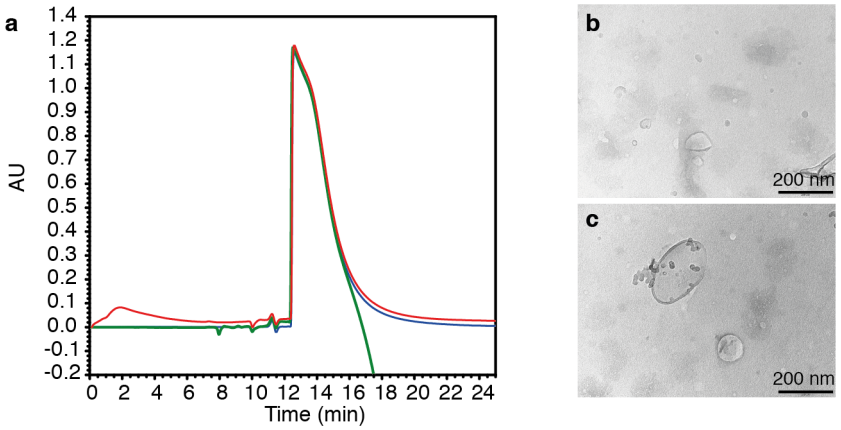

Figure 53: Aggregation of KF dilutions. Size-exclusion chromatography of 40μM KF dilutions at different waiting times: immediately after dilution (blue), after 30 min (green) and after 3 h (red). The width of the elution peak indicates peptide aggregation at all times.(**b, c**) TEM images of a 4 μM KF sample prepared by freeze-drying.

zeta potential to larger negative values, further supporting that DNA binds to the peptide aggregates preventing their growth. In general, significant differences are not observed between DLS measurements of KF stock aliquots stored at −20°C or solutions freshly prepared from lyophilized KF.

14.3.1 *Result with complementary techniques*

The DLS results are also in agreement with size-exclusion high pressure liquid chromatography (HPLC) measurements in which the formation of KF aggregates was observed since the initial time of dilution (Figure 53**a**). To obtain further insight into the structure of KF particles, transmission electron microscopy (TEM) images of cryofractured KF samples were performed and compared to peptide-free buffer samples, finding structures that resemble micelles of 100-200 nm (Figure 53**b,c**). The cryoprotectant properties of DMSO negatively affect the solvent sublimation during the freeze-drying process, inducing the formation of other structures that make difficult to image the particles. However, similar doughnut-shaped structures were already seen in a previous characterization of KF in water dilutions [254] suggesting that the structures seen in Figure 53**b,c** might be KF aggregates.

14.4 KF-DNA AFFINITY MEASUREMENTS IN THE EB REGIME

Pulling experiments of DNA in the presence of KF show a FDC pattern with force rips and hysteresis even if the pulling is performed at very low speeds (Figure 54**a**, blue). Low pulling speeds are par-

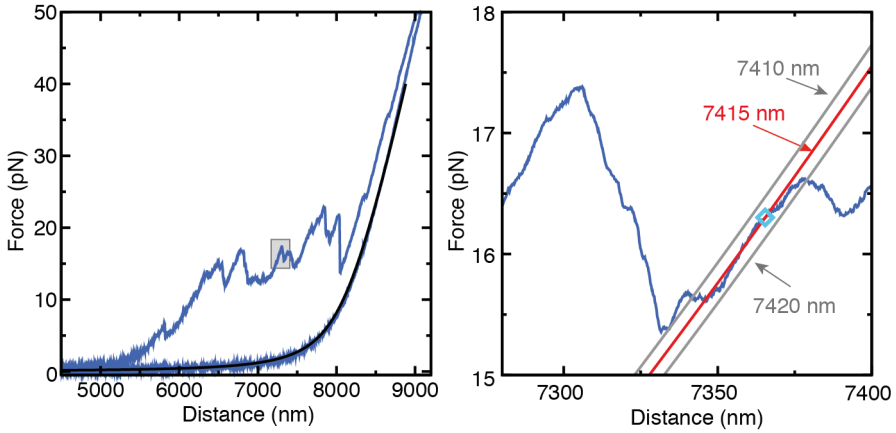

Figure 54: Statistical analysis of KF-DNA unpeeling. (a) Blue curve shows a typical FDC in a KF-DNA pulling experiment (20 Hz bandwidth, v=30 nm/s).The release cycle is fitted to the WLC model (black). The area in gray is zoomed in panel (b). **(b)** For each data point (cyan) of the FDC we determine the apparent contour length l_0 that passes closest to it according to Eq. 35 (in this particular case $l_0 = 7415$ nm, red).

ticularly useful to characterize the affinity of DNA binding to KF aggregates during the EB regime. In these experiments, the slope between two consecutive force rips reflects the elastic response of DNA with a given apparent contour length l_0. Each force rip is due to the unpeeling of a DNA segment that was electrostatically bound to the KF particle. A statistical analysis of force rips was used to determine the length of the DNA segments released during the unpeeling process [196]. This analysis is very similar to that previously used to perform single-molecule footprinting experiments (chapter 9).

For every data point (x_{exp}, f_{exp}) of a FDC, we determined its most probable apparent contour length (l_0) by finding the theoretical WLC [248] that passes closest to that point at the force f_{exp} (Figure 54**b**):

$$\left|x_{exp} - x_{WLC}\left(l_0, f_{exp}\right)\right| = \min_{l}\left(\left|x_{exp} - x_{WLC}\left(l, f_{exp}\right)\right|\right). \qquad (35)$$

The theoretical extension $(x_{WLC}\left(l_0, f_{exp}\right))$ was determined imposing the elastic parameters (l_p, S) obtained from a previous WLC fit to the relaxation curve of the same pulling experiment (Figure 54**a**, black). In this way, each experimental data point was associated to an apparent contour length l_0 (Figure 55**a**, red). A histogram of all the l_0 values showed a series of peaks that identify states that are stabilized by KF-DNA contacts (Figure 55**b**). The distance between two consecutive peaks is the length of the DNA segment released at every unpeeling event. The histogram was then fitted to a sum of Gaussians (Figure 55**c**), and the distance between the mean of consecutive peaks was calculated. The experimental distribution of unpeeling events is broad (from a few nm to ~400 nm) and follows an exponential distribution with mean size $\Delta l_0^* = 31 \pm 6$ nm (Figure 55**d**). An exponential distribution of unpeeling lengths is known to correspond to the distri-

bution of intervals expected in random partitioning a given contour length, in agreement with our hypothesis that DNA binds KF aggregates at hydrophilic spots in a random fashion.

A force vs. contour length representation (Figure 55**e**) emphasizes the release of DNA segments in a stepwise manner during the unpeeling process. To determine the mechanical work ($W = F\Delta x$) performed to disrupt each KF-DNA contact in DNA stretching experiments, we determined the average rupture force and the extension of DNA released at every unpeeling event. The average rupture force of each unpeeling event can be determined as the mean of the average force immediately before and after an unpeeling event. The rupture force distribution is broad (from just a few pN's up to the maximum forces of the pulling experiment, 50 pN) with most unpeeling events happening at forces between 10 and 20 pN (Figure 56**a**). Once the average force and unpeeling length of each event are determined, the mechanical work performed to disrupt each KF-DNA contact can be directly inferred ($W = F\Delta x$). However, this work is partially used to stretch the released DNA up to the rupture force, whereas the rest of work is dissipated into the solvent in the form of heat:

$$W_{\text{dissipated}} = \frac{F\Delta F}{k} - \Delta G_{\text{stretching}} \tag{36}$$

$$\Delta G_{\text{stretching}} = \frac{\Delta L}{L} \int_0^{x_{\text{rup}}} F_{WLC}(x)\, dx \tag{37}$$

where x_{rup} is the molecular extension of the DNA fiber at the rupture force. In the above expressions we use the fact that the FEC of the WLC model is a sole function of x/L. A histogram of the dissipated work (Figure 56**b**) shows an exponential distribution with average unpeeling energy of 13.5 ± 5 kcal/mol. This value sets an upper limit to the free energy of binding of KF to DNA.

14.4.1 *Simulation of the FEC curves*

To gain a better understanding of the affinity of DNA binding to KF particles, we used a simple theoretical model that reproduces the experimental FECs.

We considered a model previously used to characterize the DNA-dendrimer condensation transition [255]. In this model KF-DNA stretching curves are simulated using a set of N non-interacting two-state systems (Figure 57, Inset) that describe each of the contacts made between KF and DNA. Each contact can be in two conformations: formed or dissociated. The formed and dissociated states of contact i^{th} have extensions 0 and x_i respectively. Initially the N contacts are found in the formed conformation (i.e. zero extension). As the DNA

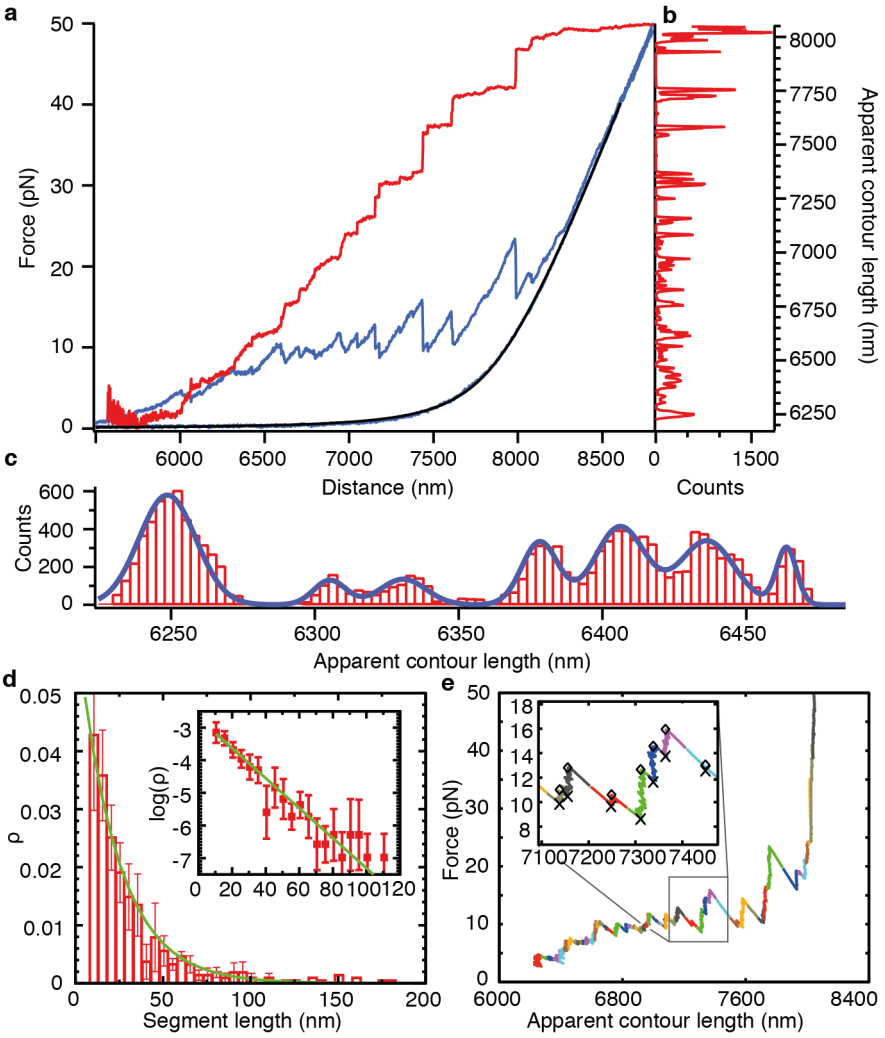

Figure 55: Analysis of DNA apparent contour length during unpeeling. (a) FDC of a KF-DNA pulling experiment (blue) and WLC fit to the release curve (black). Red curve shows the apparent contour length l_0 (right axis) for each experimental data point of the FDC. **(b)** Histogram of l_0 values. **(c)** Detailed view of the histogram (red) and fit to a sum of Gaussians (blue). **(d)** Histogram of unpeeling segment lengths Δl_0 and fit to an exponential distribution (green). Inset shows a log-normal plot (mean±SD, N=435 events, 3 molecules). **(e)** Force vs. apparent contour length representation of the pulling experiment. Each color identifies a state (apparent contour length) temporally stabilized by KF-DNA contacts during the unpeeling process. **(Inset)** Sharp transitions between states are observed. Minimum and maximum forces of every state are indicated with crosses and diamonds respectively.

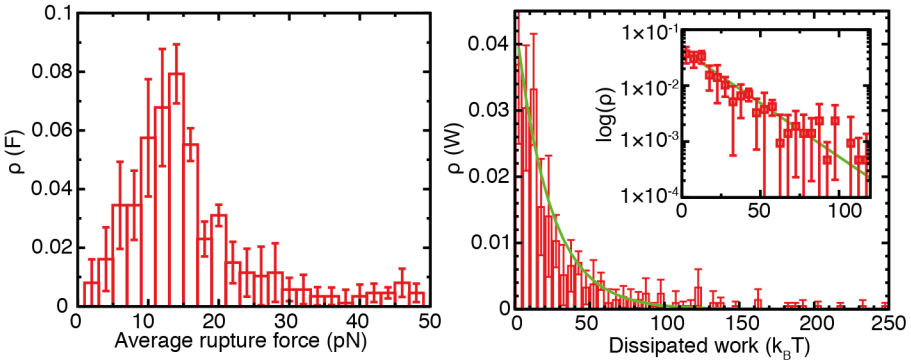

Figure 56: Rupture force and dissipated work in KF-DNA unpeeling (**a**) Average rupture force of KF-DNA contacts.(**b**) Histogram of dissipated work in individual unpeeling events (calculated using eq. 37). The histogram follows an exponential distribution of mean $23 \pm 8\ k_B T$ (green). A rightmost tail corresponding to individual unpeeling events with $W_{dissipated} \geqslant 150\ k_B T$ is observed. Inset shows a log-normal plot. Unpeeling events with dissipated work as large as $400\ k_B T$ are observed. Error bars are the statistical error measured between different molecules (N=435 events from 3 molecules for both pannels).

is stretched and force increased, the tilting of the free-energy landscape towards larger extensions favors the dissociated conformation, releasing an extension x_i. Transitions between both states are simulated as a thermally activated process characterized by a critical force f_c, and force-dependent dissociation/association rates $k_{F\to D/D\to F}(f)$:

$$k_{F\to D,i}(f) = k_a \exp\left[-\left(B_i - f x_i^\dagger / k_B T\right)\right] \tag{38a}$$

$$k_{D\to F,i}(f) = k_a \exp\left[-\left(B_i - \Delta G_i + f(x_i - x_i^\dagger)/k_B T\right)\right] \tag{38b}$$

where each segment is described by its free energy of formation ΔG_i, activation barrier B_i, extension x_i, and distance to the transition state $x_i^\dagger$. The rate k_a is a microscopic rate that does not depend on the particular contact and is equal to the attempt frequency of the molecular interaction (typically $10^5 - 10^8$ Hz). The released extension x_i is assumed to follow the experimental distribution (Figure 55**d**), and we introduced some structural disorder by assuming that ΔG and B are also exponentially distributed.

The model reproduces the essential features of the experimental curves (Figure 57) over a wide range of pulling speeds (30-500 nm/s). Despite the apparent large number of free parameters, only certain values in very specific ranges can reproduce these features[1] (see section 14.4.1.2). In brief, the experimental FECs could be well described by assuming: (a) a low binding energy of DNA to KF aggregates

1 Simulation parameters that best describe the experimental curves: $p(\Delta G) = (1/w)\, e^{-(\Delta G - \Delta G_0)/w}$ with $\Delta G \geqslant \Delta G_0 = 10\ k_B T$ and $w = 1\ k_B T$; $x_i^\dagger = 2 \pm 1$ nm; $p(B) = (1/w')\, e^{-(B - B_0)/w'}$ with $B \geqslant B_0 = 22\ k_B T$ and $w' = 5\ k_B T$. The contact-length distribution (x_i) was assumed to follow the experimental distribution shown in Fig. 55**b**: $p(x_i) = (1/w'')\, e^{-(x_i - x_{i,0})/w''}$ for $x_i \geqslant x_{i,0}$ ($p(x_i) = 0$ otherwise) with

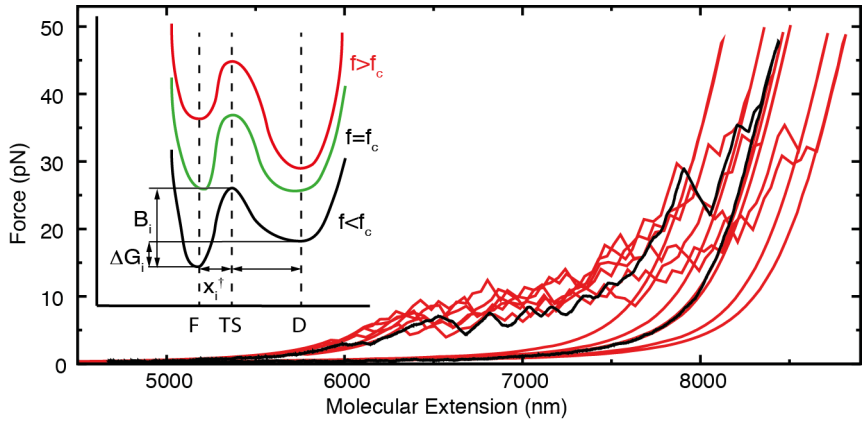

Figure 57: Modeling of DNA stretching experiments. The black curve is an experimental pulling curve after 15 min interaction with KF (v=500 nm/s). A set of six simulations of the theoretical model is shown in red. **(Inset)** Scheme of the free-energy landscape of a two-states system at different forces (f). The main parameters describing the system are: the free energy difference (ΔG_i) between the formed (F) and dissociated (D) conformations, the height of the barrier (B_i), the distance (x_i) separating the two conformations, and the distance ($x_i^\dagger$) from the transition state (TS) to the formed conformation. As the force is increased the free-energy landscape is tilted favoring the dissociated conformation above the critical force (f_c).

($\Delta G \sim 6 \pm 2$ kcal/mol); (b) a brittle unpeeling of the DNA segments ($x_i^\dagger = 2 \pm 1$ nm, the barrier lying close to the formed conformation); and (c) a broad right-tailed distribution $p(B)$ of high energy activation barriers given by $p(B) = (1/w') \exp[-(B - B_0)/w']$ with $B \geqslant B_0 = 22\ k_B T$ and $w' = 5\ k_B T$.

From these values, the critical force of the interaction f_c and the kinetic rate at the critical force $k_{F\to D}(f_c)$ were also determined:

$$f_c = \Delta G/\Delta x \sim \langle \Delta G_i \rangle / \langle x_i \rangle \sim 2 \text{ pN} \tag{39a}$$

$$k_{F\to D}(f_c) = k_a \exp\left[-\left(B - f_c x_i^\dagger / k_B T\right)\right] \sim 0.1 - 10^{-4} \text{ Hz} \tag{39b}$$

where Δx is the change in extension at f_c. Note as well, that at the critical force f_c it is satisfied that $k_{F\to D}(f_c) = k_{D\to F}(f_c)$.

14.4.1.1 *Implementation of the simulation*

The simulation is implemented as follows:

1. The molecule is partitioned between N segments where

$$N = (l_0 - l_{slack})/\langle x_i \rangle \tag{40}$$

where l_{slack} is the apparent DNA contour length when all the contacts are in the formed conformation (i.e. the minimum apparent contour length reached in a pulling experiment), and $\langle x_i \rangle$

$x_{i,0} = 8$ nm and $w'' = 24$ nm. From these parameters it can be shown that $f_c \sim 2$ pN and $k_c \sim 0.1 - 10^{-4}$ Hz. Other specific parameters for these simulations: $k_{trap} = 0.07$ pN/nm, $v = 500$ nm/s, $l_p = 35$ nm, $l_{slack} = 6500$ nm. Errors are an estimation of the range in which the features of the FDC are well reproduced by the model when each parameter is independently modified.

is the average contact length. A length x_i, free-energy of formation ΔG_i, heigh barrier B_i and distance to the transition state $x_i^\dagger$ are assigned to each segment following the aforementioned distributions:

$$\Delta G_i = \Delta G_0 - w \log(r) \tag{41a}$$
$$B_i = B_0 - w' \log(r) \tag{41b}$$
$$x_i = x_{i,0} - w'' \log(r) \tag{41c}$$
$$x_i^\dagger = 2\ \text{nm} \tag{41d}$$

where r is a random number ($0 < r < 1$). A parameter s_i is assigned to each contact to indicate whether it is formed ($s_i = 0$) or dissociated ($s_i = 1$). Initially all of the contacts are in the formed conformation. The zero-force dissociation rate constant $k_{F\to D,i}(0)$ of each contact is also determined as:

$$k_{F\to D,i}(0) = k_a \exp(-B_i/k_B T) \tag{42}$$

2. The force f at the initial position of the optical trap $x_{trap,0}$ is calculated using the WLC model (with parameters $l_0 = l_{slack}$; l_p; S; k_{trap}).

3. At every time-step dt, N contacts are randomly-picked and each of them is tested against a thermal force:

   - If a contact i is in the formed state ($s_i = 0$); a random number ($0 < r < 1$) is picked and compared to the probability of dissociating the contact due to thermal fluctuations:

   $$p_{open} = k_{F\to D,i}(0) \exp\left(f x_i^\dagger / k_B T\right) dt \tag{43}$$

   If $r < p_{open}$ the contact becomes dissociated ($s_i = 1$) and the apparent contour length of the molecule l_0 increases in $s_i x_i$.

   - If a contact i is in the dissociated state ($s_i = 1$); a random number (r) is picked and compared to the probability of forming the contact due to thermal fluctuations:

   $$p_{close} = k_{F\to D,i}(0) \exp\left[\left(\Delta G_i - f(x_i \varepsilon_{WLC} - x_i^\dagger)\right)/k_B T\right] dt \tag{44}$$

   where ε_{WLC} is an entropic correction to the molecular extension of the segment ($\varepsilon_{WLC} \approx 1 - \sqrt{k_B T/(4 l_p f)}$). If $r < p_{close}$ the contact is formed ($s_i = 0$) and the apparent contour length of the molecule l_0 decreases in $s_i x_i$.

4. After every time-step the position of the optical trap and molecular extension are updated using the pulling speed ($x_{trap,t+dt} =$

$x_{trap,t} + \nu dt$), and the force f is calculated using the WLC model (with parameters $l_0 = l_{slack} + \sum_i^N s_i x_i$; l_p; S; k_{trap}).

5. This procedure is iterated until reaching a maximum force of 50 pN; and then an equivalent time-reversed protocol is simulated (i.e. relaxation curve) by reverting the pulling speed ν.

14.4.1.2 *Verification of optimal values of the simulation*

The parameters that best reproduce the experimental FECs are described earlier in section 14.4.1. Here we present a brief description of how modifications in these parameters affect the simulated FECs, and the overall match between experiments and simulations. As the experimental FECs are of highly stochastic nature, this approach has been most useful to determine the uncertainty on the optimal parameters. A set of figures in which a different parameter is modified in each panel (the others being kept at the optimal value) is presented in Figure 58. In each panel, the optimal simulation is presented in black, whereas simulations with varying values of the modified parameters are presented in colors.

In a typical KF-DNA pulling experiment the DNA molecule can be repeatedly pulled between a minimum extension and a maximum force, and the characteristic sawtooth pattern remains visible for several consecutive pulling cycles. However, in simulations where $\Delta G_0 \leqslant 5\ k_B T$ most of the contacts remain in the dissociated conformation (60% for $\Delta G_0 = 5\ k_B T$) at the end of a simulated pulling cycle (minimum extension of $\sim 4.5\ \mu m$). This high fraction of dissociated contacts cannot explain the sawtooth pattern observed in the FEC in subsequent pulls. This suggests that a value of $\Delta G_0 \leqslant 5\ k_B T$ is too low to reproduce the experimental curves. On the other hand, for $\Delta G_0 \sim 10\ k_B T$ more than 90% of the contacts are formed again at the end of a pulling cycle (minimum extension of $\sim 4.5\ \mu m$), in agreement with the experimental results both at low (50 nm/s) and high pulling speeds (500 nm/s). Finally, for free energies greater than $15\ k_B T$ a sawtooth pattern is also observed in the relaxation curve, in disagreement with the releasing part of the observed experimental FEC for the same range of pulling speeds (Figure 58**a**).

Experimental FECs show a broad distribution of rupture forces with most rupture events occurring at forces lower than 20 pN, but with a large rightmost tail (Figure 56**a**). This phenomenology can be well reproduced by considering: (i) that the transition state is close to the formed conformation (Figure 58**b**); and (ii) that the process is characterized by a disordered ensemble of barriers rather than a single valued barrier. This structural disorder is introduced in the form of an exponential distribution of barriers with a rightmost exponential tail of width $w' \geqslant 1\ k_B T$ (Figure 58**c**).

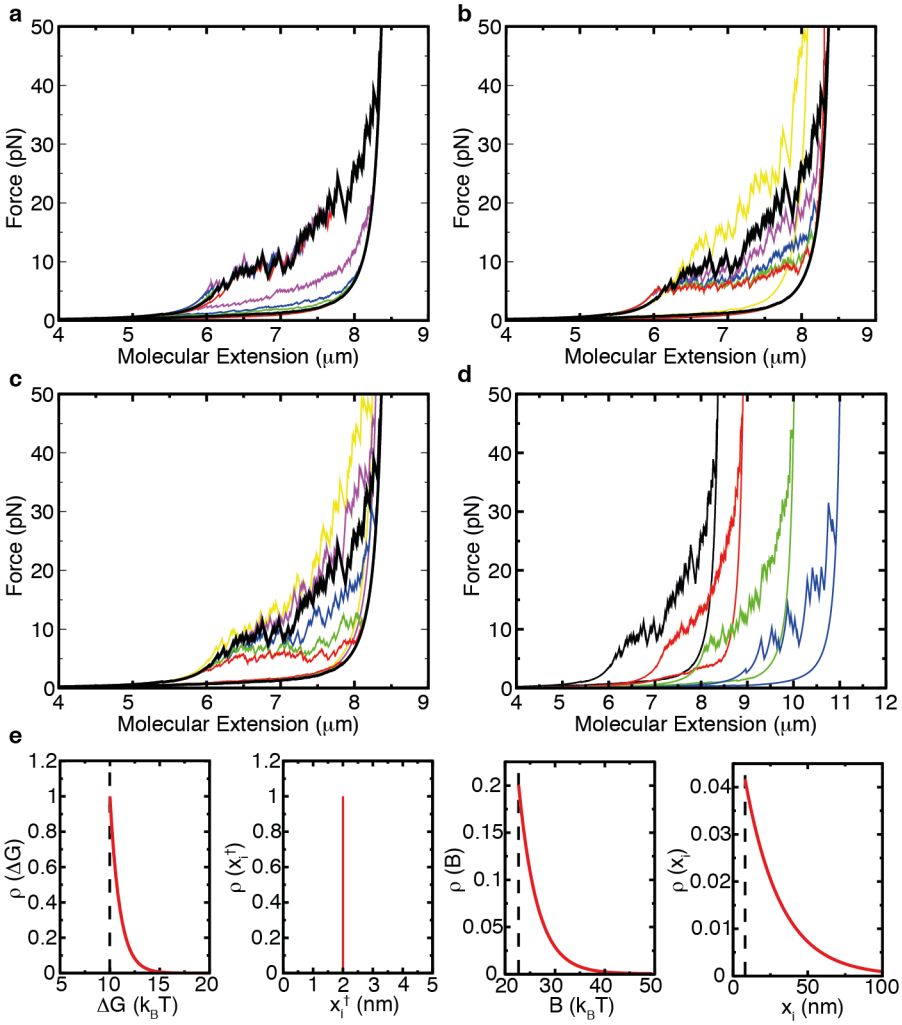

Figure 58: Simulations of KF-DNA stretching experiments. (**a**) Simulations varying ΔG_0 (minimum value of the exponential distribution of ΔG values, width $w = 1\ k_B T$): ΔG_0=5 $k_B T$ (red), ΔG_0=10 $k_B T$ (black, optimal value), ΔG_0=15 $k_B T$ (green), ΔG_0=20 $k_B T$ (blue), ΔG_0=40 $k_B T$ (purple). (**b**) Simulations varying $x_i^\dagger$: $x_i^\dagger$=10 nm (red), $x_i^\dagger$=8 nm (green), $x_i^\dagger$=5 nm (blue), $x_i^\dagger$=3 nm (purple), $x_i^\dagger$=2 nm (black, optimal value), $x_i^\dagger$=1 nm (yellow). (**c**) Simulations varying w' (exponential tail of the barrier B): w'=0 $k_B T$ (red), w'=1 $k_B T$ (green), w'=3 $k_B T$ (blue), w'=5 $k_B T$ (black, optimal value), w'=7 $k_B T$ (purple), w'=10 $k_B T$ (yellow) (**d**) Simulations varying x_i (contact-length distribution): experimental distribution $p(x_i) = (1/w'') \exp\left[-(x_i - x_{i,0})/w''\right]$ for $x_i \geqslant x_{i,0}$ ($p(x_i)$=0 otherwise) with $x_{i,0}$=8 nm and w''=24 nm (black) is compared to a single-valued contact-length: x_i= 10 nm (red), x_i=30 nm (green), x_i=100 nm (blue). Simulations are shifted by 1 μm for clarity. (**e**) Optimal distributions for the different parameters of the simulation (corresponding to the black lines in panels a-d).

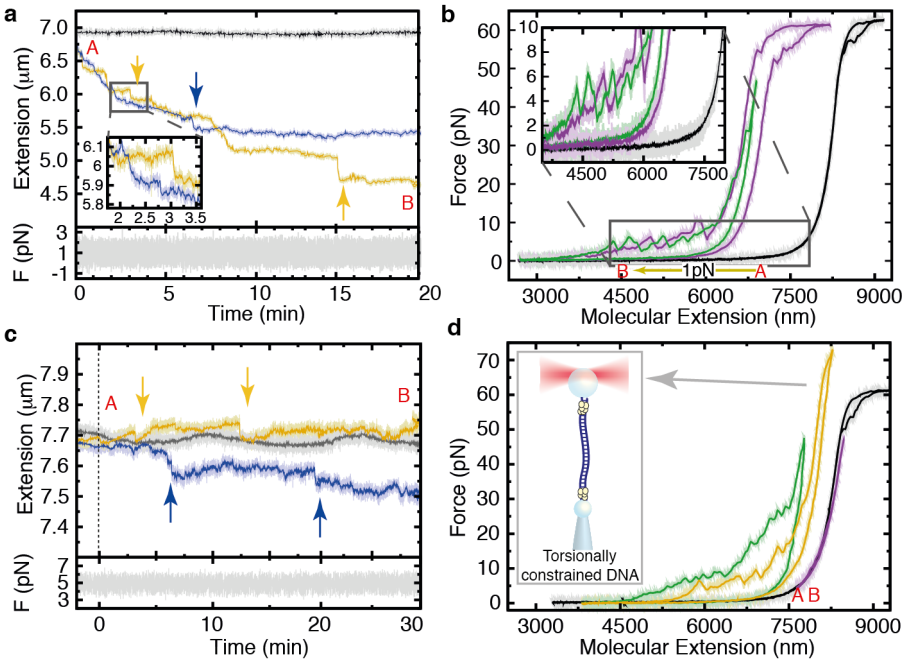

Figure 59: Kinetics of DNA binding to KF particles at a constant force. (a) DNA compaction at 1 pN. A control without peptide (gray) and two equivalent experiments at 40 μM KF (blue and yellow) are shown. The molecule was maintained at an initial extension A that relaxed at constant force down to a final value B. **(b)** Stretching of a DNA molecule before (black) and after (green, purple) the constant-force experiment at 1 pN. **(c)** KF does not compact DNA at 5 pN. A DNA molecule subjected at 5 pN is rinsed with KF and changes in the molecular extension are monitored. A control experiment without peptide (gray), and two independent experiments at 40 μM KF (blue and yellow) are shown. Large fluctuations indicative of individual binding events are observed (arrows) **(d)** Stretching of a DNA molecule only at forces higher than 5 pN immediately after the constant-force experiment at 5pN (purple). If the force was relaxed below 5 pN (green and yellow), the characteristic sawtooth pattern was immediately recovered. The FEC of that molecule before flowing KF is shown in black. For all plots, raw data (1 kHz) is shown in light colors, and filtered data (1 Hz bandwidth for kinetic experiments, 10 Hz bandwidth for pulling experiments) is presented in dark colors. Pulling speed is 500 nm/s.

To model the experimental FECs, we have assumed that the distribution of released lengths (x_i) follows the experimental distribution shown in Figure 55**d**. In Figure 58**d** a simulation using the experimental distribution is compared to simulations in which the system is characterized by a single contact length (x_i) rather than an exponential distribution. Finally the optimal distributions for the different parameters of the simulation are shown in Figure 58**e**.

14.5 ELECTROSTATIC BINDING REGIME INVESTIGATED IN CONSTANT FORCE EXPERIMENTS

We followed the kinetics of DNA compaction by performing constant-force experiments at forces such that EB prevails (molecular extension $\geqslant$ 5.5 μm, Figure 48**b**). In these experiments, the DNA molecule was

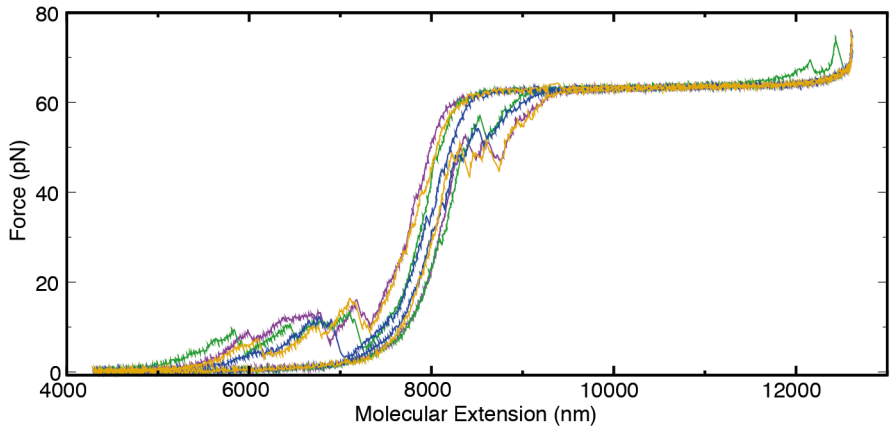

Figure 60: KF binding does not change the overstretching transition. Four consecutive stretching curves (purple, green, blue and yellow) of a 24-kb DNA molecule after incubation with 50 μM KF. The molecule is fully overstretched at each pulling cycle. The sawtooth pattern at low extensions is clearly visible at each cycle. Data was filtered at 100 Hz bandwidth, v=1000 nm/s.

maintained at constant force using a force-feedback, and we followed the time-evolution of the molecular extension while KF was flowed into the microfluidic chamber (details on the flow set-up are found in Appendix B).

At 1 pN, a fast compaction took place (Figure 59**a**). The extension was reduced up to 40% in 20 min at a reproducible rate. This compaction is characterized by intermittent drops of extension that shorten the molecule by hundreds of nanometers in a few seconds (Figure 59**a**, arrows). Pulling curves performed after this experiment (Figure 59**b**) showed again the characteristic sawtooth pattern observed in previous KF-DNA pulling experiments.

In contrast, at 5 pN the molecular extension remained almost constant within 100 nm after flowing KF for more than 30 min (Figure 59**c**). Still, intermittent large fluctuations on the order of tens of nanometers were often detected (Figure 59**c**, arrows). These large fluctuations were never observed in controls without KF (Figure 59**c**, gray), and we attribute them to individual binding events. Pulling cycles performed between 5 and 40 pN immediately after the peptide flow show a slight decrease in the persistence length and weak hysteresis effects suggesting very weak binding of KF to DNA (Figure 59**d**, purple). Only by further decreasing the extension and force of the molecule full binding events were observed (Figure 59**d**, green), recovering again the typical KF-DNA sawtooth pattern.

Interestingly an overstretching transition was not always observed (Figure 59**d**, yellow). We attribute this to the recruitment of DNA segments close to both ends of the tethered molecule by KF particles that induce a torsionally constrained fiber, inhibiting the overstretching transition (Figure 59**d**, Inset).[256] Otherwise, KF binding does not suppress or tilt the overstretching plateau as observed for DNA

intercalators.[67] Moreover, the characteristic sawtooth pattern of KF remained visible after fully overstretching the DNA (Figure 60).

CHAPTER 15

RESULTS: DNA UNZIPPING AND INTERACTION OF KAHALALIDE F WITH ssDNA

In this section we show that DNA unzipping experiments can also be used to investigate the complexation process between KF and DNA, providing complementary information to the results obtained in the stretching configuration (previous section). First, we show how the characteristic sawtooth pattern of DNA unzipping globally changes after binding of KF, reflecting the increased stability of the complex. Remarkably, DNA unzipping also provides experimental evidence that KF binds to ssDNA. To further investigate this unexpected result, we developed a method to generate a long ssDNA molecule for optical tweezers manipulation. We then performed experiments with this template finding that binding of KF to ssDNA also proceeds in a two-steps mechanism in which ssDNA compaction takes place first followed up by a rigidification of the molecule. These two steps are then related to the previously determined electrostatic and hydrophobic binding mechanisms. Finally, we perform AFM images of KF-ssDNA complexes observing the formation of aggregation spots, and finding that binding to KF favors dsDNA hybridization; a result very similar to what has been reported for aggregating amyloid peptides.

15.1 UNZIPPING EXPERIMENTS REVEAL DIFFERENT BINDING MODES OF KF TO dsDNA AND ssDNA

To perform DNA unzipping experiments we used the 6.8-kb DNA hairpin construct previously used for the DNA wash-off experiments

with Thiocoraline (chapter 8.1.1, synthesis details in Appendix D.2). A single DNA hairpin can be specifically attached between an optically-trapped antidigoxigenin coated bead and a streptavidin-coated bead held by air suction in the tip of a glass micropipette (Figure 61**a**, Inset).

In a DNA unzipping experiment, when the distance between the beads is increased, a force is applied to the dsDNA handles that is in turn transmitted to the hairpin fork. At a force of ~ 15 pN the basepairs cannot withstand this force and break releasing a certain amount of ssDNA bases. This increases the total length of the molecule, causing a drop in the force exerted by the optical trap. When the distance between beads is increased again, the force increases and this same process is repeated for several cycles. This gives rise to a sawtooth pattern (Figure 61**a**, gray) in which force rips are related to the base-pairing stability (DNA sequence), and gentle slopes to the elastic response of the ssDNA generated during the unzipping process [62]. Consequently, the unzipping pattern of a DNA molecule is a fingerprint of its base-sequence, and changes of that pattern after flowing a ligand or KF indicate a direct interaction between the peptide and DNA. Moreover, with this set-up we can explore the effect of KF on a DNA molecule maintained at zero force and forming a random-coil (the force stretches the linkers but not the hairpin). This is particularly useful to investigate the KF-DNA interaction as it provides a set-up in which binding to DNA and the subsequent compaction takes place in a similar fashion as it would for DNA in a bulk solution (e.g. as in the molecules used for AFM imaging).

To perform the experiments a hairpin was tethered and partially unzipped, maintaining at least half of the dsDNA stem open (Figure 61**a**, dashed line) and then KF was flowed into the chamber. In this configuration the released ssDNA is long but rigid enough to severely restrict thermal fluctuations in the molecular extension ($\sqrt{\langle \delta x^2 \rangle} \sim 20$ nm).

Hence an advantage of this set-up is also that the long separation between the hairpin and the polystyrene beads (~ 4 μm) inhibits any interaction between the beads and both the dsDNA region and linkers.

After flowing KF for 3 min, the unzipping pattern substantially changed (Figure 61**a**) and forces up to 22 pN were needed to unzip the DNA. We attribute this to the increased force required to simultaneously break the base-pairing interactions and unpeel DNA segments from the KF particles. Consecutive unzipping curves show that ssDNA remains bound to KF particles at the maximum forces (25 pN). Surprisingly enough, the re-zipping trajectories overlapped with the re-zipping curves of naked DNA over a wide range of extensions ($\geqslant$1500 nm). This is in agreement with a re-annealing mechanism in which re-hybridization takes place first, followed by the formation of

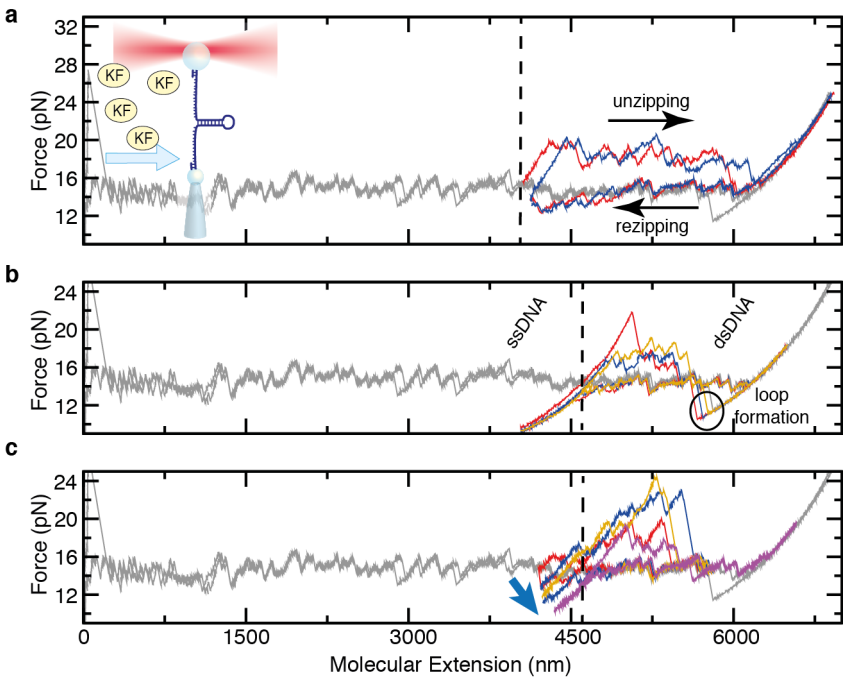

Figure 61: Unzipping experiments show that KF binds both dsDNA and ssDNA. **(a)** An unzipping pattern of the DNA hairpin before incubation with KF is fully represented in gray in each panel. The dashed line represents the position at which the molecule remained unzipped during the peptide flow. Consecutive pulling cycles of the dsDNA stem region after incubation with KF (blue, red) show a strong distortion of the unzipping pattern. However, the re-zipping of the hairpin remains unaffected indicating that the molecule can hybridize again. **(Inset)** A 6.8-kb hairpin is maintained partially unzipped leaving less than half of the dsDNA stem closed during the incubation with KF. This configuration prevents KF-DNA interactions potentially mediated by the beads. **(b)** The region of the hairpin that remained as ssDNA during the peptide flow (left of dashed line) cannot hybridize again in contrast to what happens without KF (gray curve) or in the region maintained as dsDNA during the peptide flow (right of dashed line). Three pulling cycles are plotted in blue, red and yellow. **(c)** If the hairpin is rinsed with KF and then submitted to several consecutive pulling cycles, the non-hybridizing region increases with time. Four pulling cycles that reflect this trend are shown (red, blue, yellow and purple). Data is filtered at 10 Hz bandwidth, v=50 nm/s.

the complex. Similar results have been observed in the formation of amyloid nucleic acid fibers, in which the binding of amyloid peptides to oligonucleotides promotes their hybridization [107].

Note that in this experiment we only unzipped the region of the hairpin that remained in double-stranded form while KF is flowed (right of the dashed line in Figure 61**b**). However, when we tried to rezip the region of the hairpin that remained as ssDNA during the peptide flow (left of the dashed line in Figure 61**b**), we could not recover the characteristic unzipping pattern of the molecule. This indicates that KF can bind ssDNA in a way that prevents re-hybridization of ssDNA strands. If the molecule was continuously submitted to unzipping/re-zipping cycles, the region of the DNA hairpin that previously re-zipped progressively looses that capability (Figure 61**c**) suggesting that KF is slowly binding to the stretched ssDNA (arrow in Figure 61**c**).

These results suggest that KF binds to the phosphate backbone in a configuration that does not interfere with base-pairing interactions when DNA is in its double-stranded form. However, when DNA is in its single-stranded form, KF can adopt configurations that interfere with the re-zipping of the molecule. These experiments also show that the interaction of the peptide with ssDNA is slow (in the order of minutes), as only the ssDNA regions of the hairpin that remained exposed for long times to the peptide were unable to re-hybridize.

15.2 INTERACTION OF KF WITH ssDNA

15.2.1 *Preparing a long ssDNA molecule for optical tweezers manipulation*

To characterize the interaction of ssDNA with KF aggregates we developed a simple method to generate a long ssDNA template (13-kb) for optical tweezers manipulation (Figure 62**a**). First, a 6.8-kb DNA hairpin is tethered with the optical tweezers and mechanically unzipped. Once the hairpin is fully unzipped, a solution containing a 30-base oligonucleotide complementary to the loop and its flanking region is flushed into the microfluidics chamber (sequence details in Appendix D.2). The binding of the oligo to the stretched ssDNA prevents the re-zipping of the hairpin when the tension is released. In fact, the entropic effect to close 30 bases within a loop and the large bending rigidity of the dsDNA segment in the loop region creates a large kinetic barrier that must be overcome thereby inhibiting the refolding of the native structure at forces much lower than the average unzipping force (15 pN). By using this setup we could therefore measure the elastic response of ssDNA down to forces lower than 1 pN (Figure 62**a**). It is important to take into account though, that the blocking oligonucleotide can also nonspecifically bind to other

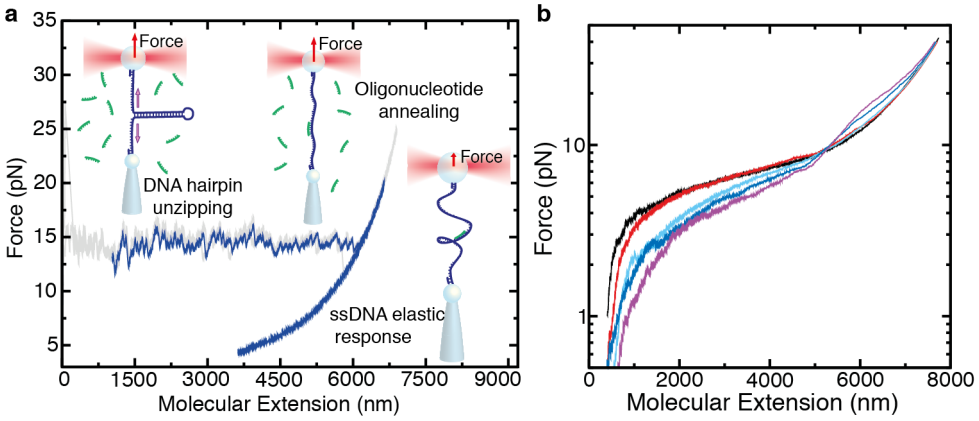

Figure 62: Method used to generate a long ssDNA molecule. (a) The specific binding of an oligonucleotide to the hairpin inhibits the hybridization of the molecule (blue) at forces lower than the average unzipping force. The full force-extension curve of the hairpin is plotted as a reference (gray). **(b)** Effect of the blocking oligo concentration in the elastic response of ssDNA due to nonspecific binding along the molecule. "Flushing oligo" method (black) compared to a constant concentration of 1 nM (red), 10 nM (cyan), 100 nM (blue) and 1000 nM (purple) blocking oligo. Experiment at 1 M NaCl.

regions of the hairpin that are partially complementary to it. This binding is not desired as it creates short segments of dsDNA that modify the elastic response of the ssDNA and interfere with the formation of random nonspecific secondary structures along the tethered ssDNA at low forces (Figure 62**b**). This effect is strongest at high ionic strengths due to the increased stability of duplex DNA. A way to minimize nonspecific binding is to use a very low concentration of blocking oligo, however this also increases the time required to obtain a ssDNA tether. Consequently, we performed the experiments using the "flushing oligo" approach. Briefly, once a DNA hairpin was tethered and fully unzipped, we flushed a relatively high concentration of oligo into the microfluidics chamber (50-250 nM depending on the ionic condition)[1]. In this way, the binding of the oligo to its complementary sequence in the hairpin took place in a few seconds. Then the DNA hairpin was briefly rinsed again with oligo free buffer to remove oligos nonspecifically bound to other regions of the hairpin that are not fully complementary. By doing so, it was possible to rapidly and reliably obtain ssDNA (Figure 62**b**).

This methodology allowed us to study in detail the interaction of KF with ssDNA (detailed in the following section), as well as the elastic response of ssDNA and the formation of non-specific secondary structure in ssDNA over a wide range of forces and salt conditions (see Appendix A.3). As a future perspective this method should be very useful to investigate the kinetics and thermodynamics of non-specific secondary structure formation (e.g. investigating ssDNA

1 The oligo solution was sometimes incubated at 45°C for 10 min to remove competing secondary structures created after repeated freeze-thaw cycles.

molecules of varying length and base-pair composition) as well as its interaction with single-stranded DNA binding (SSB) protein.

15.2.2 *Kinetics of binding of KF to ssDNA*

To investigate the kinetics of binding of KF to ssDNA, first a ssDNA molecule was tethered as previously described, and its elastic properties measured with the WLC and FJC models (WLC: $l_p = 0.87 \pm 0.04$ nm, $l_c = 0.69 \pm 0.01$ nm/bp; FJC: $l_k = 1.55 \pm 0.04$ nm, $l_c = 0.57 \pm 0.01$ nm/bp). These results being in agreement with previous measurements in which alternative methods were used to generate the ssDNA [249, 62]. A detailed analysis and results at other ionic conditions are found in Appendix A.3.

We then flowed KF into the microfluidics chamber (details on the flow set-up in Appendix B) and followed the molecular extension at a constant force using force-feedback. At 5 pN we observed a slow compaction (~20-30 min) with an absolute reduction in extension close to 16% (Figure 63**a**, upper panel), demonstrating that ssDNA binds KF. This effect was reproducible within different experiments, and the slow kinetics agree with the results from unzipping experiments (Figure 61**c**). We also measured the time evolution of the stiffness of the KF-ssDNA fiber by recording the magnitude of the thermally induced fluctuations in the molecular extension (details in section 15.2.3). At 5 pN the molecule stiffened with time at a rate of $(13 \pm 5) \cdot 10^{-3}$ pN/(nm·s) (Figure 63**a**, middle panel). However this change was only observed 10-15 min after compaction of the fiber started, suggesting that stiffness changes are mostly due to the hydrophobic collapse of the KF-ssDNA aggregate rather than electrostatic binding of ssDNA to KF particles.

Pulling curves obtained after the peptide flow (Figure 63**b**) also show force rips in the stretching curves. However the sawtooth pattern was smoother than for dsDNA suggesting the occurrence of fewer events and higher unpeeling forces (Figure 63**b**). We attribute this to the increased hydrophobic forces that stabilize the KF-ssDNA complex, that also lead to a systematic shortening of the effective contour length of the ssDNA. At a higher stretching force of 10 pN, KF did not induce compaction of ssDNA though, but intermittent jumps in the extension were observed, indicative of individual binding events. However, KF-ssDNA compaction could be induced by lowering the force down to 5 pN (Figure 63**c**) reproducing the results reported in Figure 63**a**, and confirming that KF compacts ssDNA at 5 pN but not at 10 pN. At this higher force moreover, the stiffness of the molecule remained constant within the resolution of measurements (Figure 63**d**).

Binding of ssDNA to KF aggregates was further corroborated by AFM. ssDNA molecules (2743-bp) were generated by heat denatura-

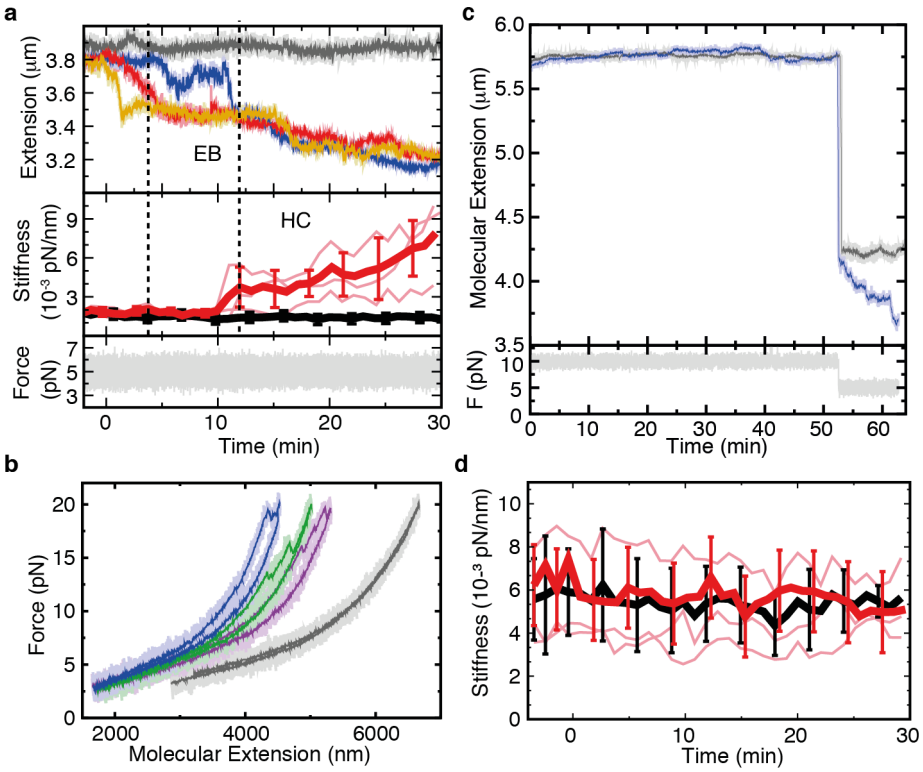

Figure 63: Kinetics of ssDNA binding to KF particles. (a) Compaction and stiffening of ssDNA at 5 pN are representative of EB and HC respectively. (Top) Extension of a ssDNA molecule rinsed with KF at 5 pN. Three independent experiments at 40 μM KF (blue, red and yellow) and a control without peptide (gray) are shown (compaction starts at t=0). (Middle) Average stiffness of the ssDNA molecule at 5 pN during the peptide flow (red) and a control without peptide (black). Three individual experiments are shown in light red. The stiffness is measured from the fluctuations in the trap position. **(b)** Stretching curves of a ssDNA molecule before (gray) and after incubation with 40 μM KF at 3 pN for 25 min (purple, green, blue). (**c**) A ssDNA molecule is maintained for more than 50 min at a constant force of 10 pN with a flow of KF, without observing a significant decrease in extension (blue). However, when the force is lowered to 5 pN (t = 52 min) a compaction equivalent to that reported in panel (a) is seen. A negative control without peptide in the flowed buffer does not show DNA compaction (gray). (**d**) Average stiffness of ssDNA at 10 pN during the first 30 min of the peptide flow (red) compared to a negative control without peptide (black). Three individual experiments are shown in light red. In contrast to the results shown in panel (a), the molecular stiffness remains constant in time.

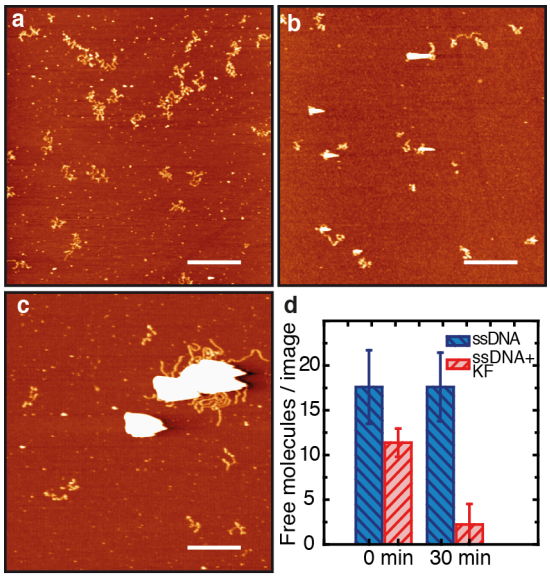

Figure 64: AFM images of KF-ssDNA complexes. (a) ssDNA molecules (1.8 nM molecules, 5 μM nucleotides) are adsorbed on a mica surface showing a much compact conformation than dsDNA due to its lower persistence length. **(b)** Immediately after mixing KF (100 μM) with ssDNA molecules (5 μM nucleotides) we observe the formation of aggregation spots and a substantial decrease in the number of ssDNA molecules per image. **(c)** After 30 min incubation at room temperature, these effects are more evident as big aggregates are seen. **(d)** Histogram of ssDNA molecules at different incubation times (0 and 30 min) with and without KF, determined as the average number of free individual molecules per image of 4 μm^2 (mean±SD, N⩾6).

tion and fast cooling down to 4°C (Figure 64**a**). Addition of KF to the ssDNA preparation triggered the formation of aggregation spots immediately after mixing (Figure 64**b**). Interestingly enough, longer incubation times yielded a reduction of free ssDNA molecules, together with the formation of large aggregates of KF surrounded by double stranded DNA (Figure 64**c-d**). This observation suggests that KF induces the rehybridization of ssDNA, as ssDNA molecules do not anneal after 30 min incubation at room temperature in the absence of KF. This is in agreement with the unzipping results (section 15.1), and is likely due to the fact that KF traps and maintains close in space different ssDNA molecules that are occasionally able to rehybridize.

15.2.3 *Measuring tether stiffness from distance fluctuations*

We have seen that KF induces an increase of the molecular stiffness of ssDNA after long incubation times. In this section we present a brief explanation on how the molecular stiffness is determined from the width of the thermal fluctuations of the optically trapped bead in the presence of a force-feedback.

A bead confined in an optical trap fluctuates around its equilibrium position due to thermal fluctuations. By the equipartition theorem the

fluctuations of the bead position along the stretching direction are directly related to the effective stiffness of the system formed by the tethered molecule and the optical trap [257]:

$$< \delta y^2 >=< y^2 > - < y >^2 = \frac{k_B T}{k_{trap} + k_{mol}} \quad . \tag{45}$$

In the KF-ssDNA experiments performed at 5 pN, the width of the fluctuations of the position of the optical trap decreases with time (Figure 65**a**). This is in agreement with an increase of the molecular stiffness as derived from Eq. 45.

Indeed, in an ideal force-feedback the stiffness of the trap vanishes and the bead fluctuations are only determined by the stiffness of the tether:

$$< \delta y^2 >= \frac{k_B T}{k_{mol}} \quad . \tag{46}$$

In this case the fluctuations of the position of the optical trap are expected to match those of the bead ($< \delta y^2 >$). However, in our experimental set-up a finite frequency feedback of 1 kHz is used (*i.e.* the force-feedback corrects the position of the optical trap by moving the piezoelectric actuators at a 1 kHz rate), and Eq. 46 is not satisfied. To extract the value of k_{mol} we followed a phenomenological approach that uses a modified version of Eq. 46 containing a proportionality constant c:

$$< \delta y^2 >= c\frac{k_B T}{k_{mol}} \quad . \tag{47}$$

The constant c includes all effects of the finite frequency of the feedback. The value of c has been obtained by fitting a set of measurements of ssDNA and dsDNA tethers at different average forces. The fluctuations in the position of the optical trap remain inversely proportional to k_{mol} in the investigated range of stiffness ($1 - 10 \cdot 10^{-3}$ pN/nm) (Figure 65**b**), verifying the validity of the method based on Eq. 47. This calibration method has been used to estimate the changes in stiffness that KF induces in ssDNA. The stiffness of the molecule during the peptide flow is measured in short time-windows to ensure reliable measurements of $< \delta y^2 >$.

Data is recorded at 1 kHz and low frequencies are filtered out to remove instrumental drift as experiments run for long times (30 min). Filtering low frequencies is also important to correct the changes in extension due to the compaction of the DNA with KF. A time-window of 3 s has been found to be the optimal value to reduce drift without affecting the measurements (vertical dashed line in Figure 65**c**). Using a shorter time-window removes fluctuations that are relevant to

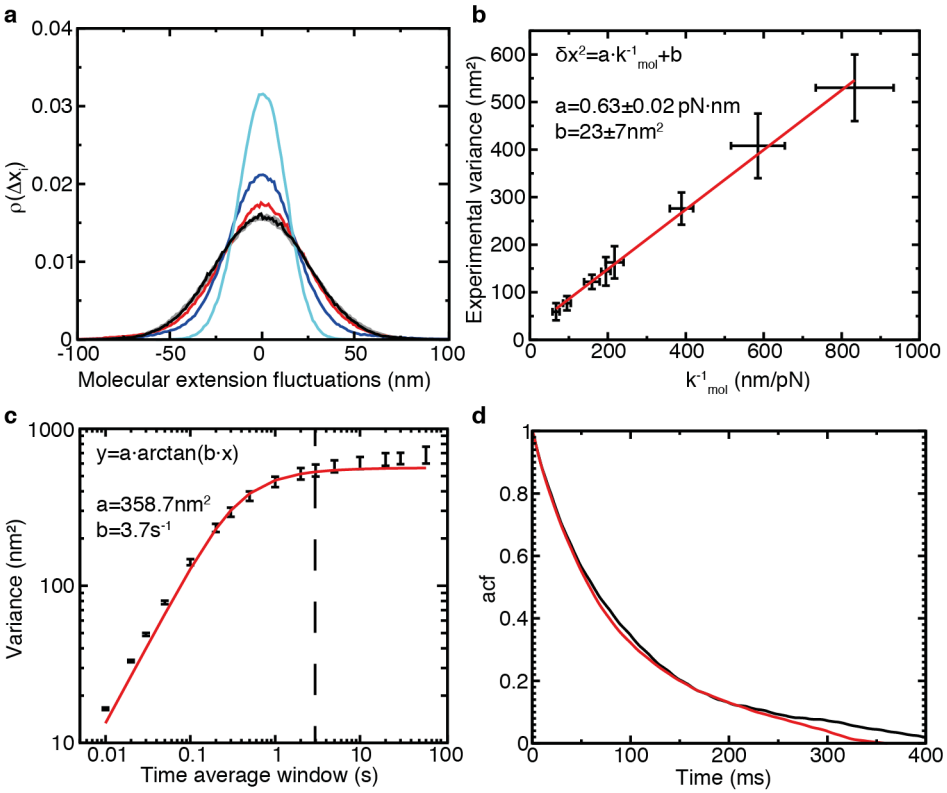

Figure 65: Stiffness determination from trap-position fluctuations. (**a**) In a constant-force experiment the molecular extension fluctuates around its equilibrium position ($\Delta x_i = x_i - \overline{x}$) due to thermal fluctuations. Fluctuations in extension of a tethered naked ssDNA molecule at 5 pN (black) and after flowing KF for 10 min (red), 20 min (blue), and 30 min (cyan). Fluctuations show a decreasing width, and hence an increasing stiffness. A negative control without KF is plotted in dark gray and overlaps with naked ssDNA. (**b**) Phenomenological calibration of the molecular stiffness. The experimental variance of the trap position for a set of ssDNA and dsDNA molecules at different average forces was measured with the force-feedback protocol. The stiffness of the molecule (k_{mol}) was determined as the derivative of the force-extension curve at each force. (**c**) Drift correction of the experimental data. The position of the trap is recorded at 1 kHz, and drift is removed by subtracting the average position using time-windows in the range 0.01-60 s (black). A window of 3 s (dashed line) is the optimal value to remove drift without distorting the measurement of the fluctuations. Data is fitted to an arctangent function (red). Measurements correspond to a 13-kb ssDNA molecule at 5 pN (N=5). (**d**) Autocorrelation function of the distance during two constant force experiments at 8 pN for a ssDNA molecule after subtracting the drift (baseline). Data with (red) and without (black) a buffer flow are shown.

determine the stiffness of the tether due to the large autocorrelation time shown by the data (Figure 65**d**) presumably caused by the force-feedback.

CHAPTER 16

DISCUSSION OF THE RESULTS

By combining single molecule techniques and bulk measurements we showed that KF forms particles that bind and compact DNA. Our measurements reveal that this process is characterized by two distinct phases controlled by the molecular extension of the DNA. First, there is a fast and weak binding regime determined by electrostatic binding (EB) to positive residues exposed on the surface of the KF particles (Figure 66**a**). This binding is triggered by spontaneous bending fluctuations along DNA. Upon reduction of the molecular extension a slow remodeling of the KF-DNA complex takes place; we propose that this new regime is led by the formation of new hydrophobic contacts that stabilize a hydrophobically collapsed (HC) structure.

A capture distance separating both regimes is identified, corresponding to a relative extension of 66%. Remarkably enough, theoretical studies between spherical charged aggregates and oppositely charged polymeric chains predicted a capture distance leading to the irreversible adsorption of the chain to the aggregates [258]. The recruitment of DNA segments along the surface of the aggregate (EB) is driven by the electrostatic attraction between the negative charge of DNA and the positive charge of L-Orn residues that are most likely exposed on the surface of the particle forming hydrophilic spots (Figure 66**b**). This interpretation is supported by the following facts: (i) The persistence length of DNA is reduced during the initial binding of the peptide, suggesting a charge compensation that reduces self-repulsion of the DNA phosphate backbone (ii) A KF analog without a positive charge does not bind to DNA (Figure 49**f**), providing

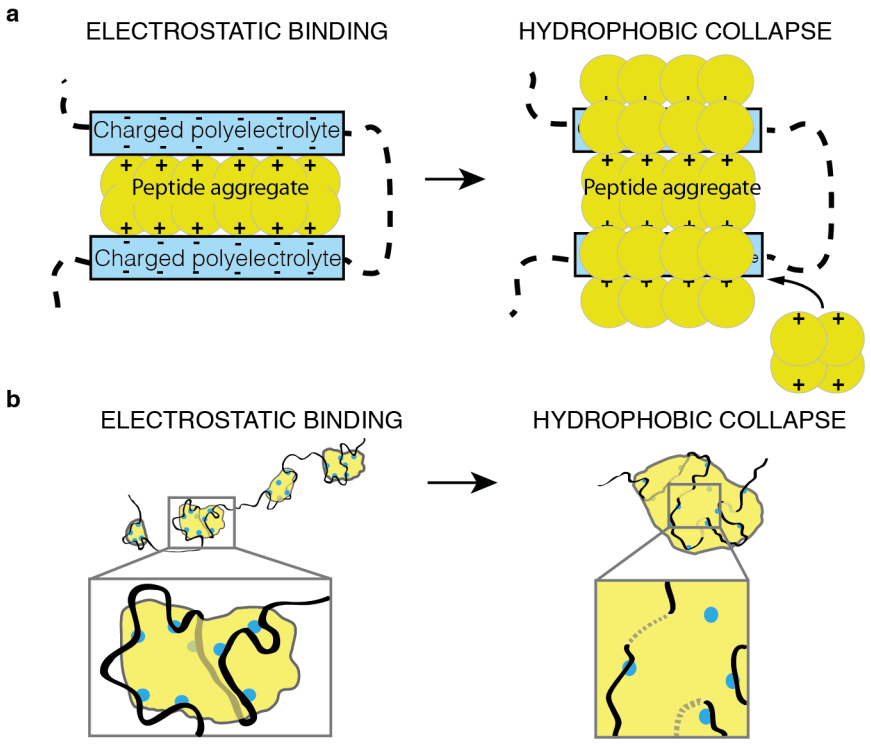

Figure 66: Model of KF-DNA complex formation. (**a**) Scheme of the two kinetic steps during the formation of the KF-DNA complex: Electrostatic binding (EB) and Hydrophobic collapse (HC). At low forces DNA (blue) binds to KF particles (yellow) due to the electrostatic attraction of the phosphate backbone to positive residues exposed on the aggregate surface (EB). This process is followed by the formation of new hydrophobic contacts between aggregates that form a larger collapsed structure (HC). (**b**) Pictorial representation of how bending fluctuations induce the EB of DNA (black) to electrophilic spots (blue) on the aggregate surface, and how hydrophobic interactions between peptides drive to a collapsed structure in which DNA becomes entrapped within the aggregated complex.

evidence that the positive charge is essential for binding (iii) A salt titration shows that binding of KF to DNA is inhibited at high salt condition, and that the strength of the interaction increases with decreasing ionic strength (Figure 49) (iv) The zeta potential value of KF, DNA, and KF-DNA complexes (Table 2) indicate different surface charge densities depending on whether DNA is complexed with KF or not. In addition, the phenomenology observed during the EB regime cannot be understood from the isolated action of KF peptides (each bearing a single positive charge). It must be due instead to the concurrent action of several positive charges that are contained in each peptide aggregate, which are then able to electrostatically interact with DNA in a similar fashion that dendrimers or other polycationic agents do. Modeling of the experimental results shows that EB of DNA to KF particles is consistent with an exponential distribution of unpeeling lengths (average value $\Delta l_0 = 31 \pm 6$ nm), a low binding energy ($\Delta G \sim 6 \pm 2$ kcal/mol), a brittle unpeeling process ($x_i^\dagger = 2 \pm 1$ nm), and a disordered collection of activation barriers

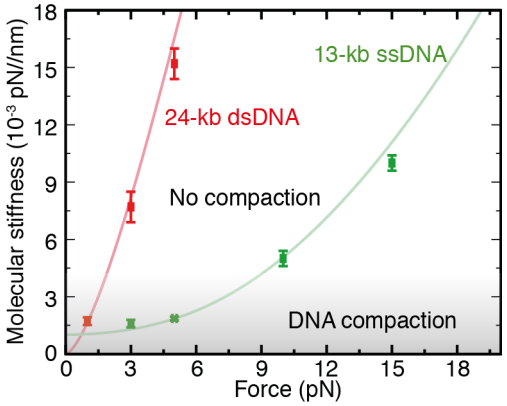

Figure 67: Phase diagram of KF-DNA compaction Phase diagram showing conditions that trigger DNA compaction due to EB of KF in constant force experiments. Average stiffness of the dsDNA (red) and ssDNA (green) molecules before flowing KF in the constant force experiments ($2 \leqslant N \leqslant 9$, mean±SD). We only observe DNA compaction in the region highlighted in gray, suggesting that compaction depends directly on DNA bending fluctuations *via* template stiffness. A WLC model (red line) and FJC model (green line) with the parameters determined in the main text for dsDNA and ssDNA respectively are plotted as a reference.

higher than B_0 =22 k_BT with an exponential right tail of width $w' = 5$ k_BT. These values are comparable to those obtained for nucleosomal particles that show a single characteristic unpeeling length of 26 nm and a single energy barrier of 36-38 k_BT [259]. The force at which we observe dsDNA compaction (1pN) is in the same range of forces that has been reported for other DNA compacting molecules such as histone-like FIS and HU proteins [260, 261] or polycationic condensing agents [255, 262, 263] .

Unzipping experiments show that KF also forms complexes with ssDNA. Interestingly, unzipping experiments together with AFM show that KF promotes dsDNA hybridization. These results agree with previous studies on the formation of amyloid nucleic acid fibers, which showed how charged surfaces of peptide complexes recruit oligonucleotides and promote their hybridization [107]. The two distinct phases observed in the formation of KF-dsDNA complexes (EB and HC) were also seen for ssDNA, showing that the mechanism of complexation is similar in both polyelectrolytes. Yet, our experiments suggest a stronger stabilization of the complex for ssDNA, which is likely due to its increased hydrophobicity and lower persistence length. Indeed, we have found that KF can compact ssDNA at higher forces than dsDNA does, indicating the role of spontaneous bending fluctuations to initiate EB. Finally, constant-force experiments with both DNA substrates show a correlation between the degree of compaction of the molecule and the effective rigidity of the tether (Figure 67).

In relation to its biological activity, and as previously showed for amyloid fibrils [129, 130, 220], the enhanced positive charge of KF aggregates could allow them to interact more effectively with other

polyanionic molecules such as polysaccharides (*e.g.* glycosaminoglycans). Whether or not the cytotoxic effects of KF are related to its interaction with polyanions, we hypothesize that these interactions could play a role in modulating the activity of the peptide either intracellularly or in the extracellular matrix. In particular KF might also interact with the phospholipids of the plasma membrane inducing the formation of pores and cell necrosis (see section 13.2.2) [236].

CHAPTER 17

CONCLUSION

The work shown in this part of the thesis represents the first attempt to extract quantitative information about the binding affinity and kinetic steps involved in the interaction between a nucleic acid (DNA) and an anticancer self-aggregating peptide (KF) at the single molecule level [182]. To date most studies of aggregation kinetics have been performed using ensemble techniques where the individual behavior of molecules cannot be distinguished. Using optical tweezers, we have shown that KF binds DNA in two kinetic steps (an initial electrostatic binding that is followed by an hydrophobic collapse of the peptide-DNA complex), and characterized the spectrum of binding affinities, kinetics barriers and lengths of DNA segments sequestered within the KF-DNA complex. The proposed methodology is not limited to the characterization of amorphous aggregates [120]. Protein aggregation, a topic of major interest due to the role of the aggregation of misfolded proteins in neurodegenerative diseases [264], might be well addressed using single molecule force spectroscopy. As well, AFM images of the nucleoid-associated proteins Dps [265, 266] and of the drug cisplatin [267] are very similar to those we found here for KF-DNA, suggesting that many biochemical studies of protein complexes that face similar aggregation or compaction phenomena are susceptible of being studied with this approach. For instance, the aggregation of the splicing factor MBNL1 by mutant mRNA hairpins is at the core of Myotonic dystrophy type I [268, 269]. Research on the formation of these RNA-protein aggregates, and of peptides that disrupt this interaction [270] could greatly benefit from the aforementioned

approaches. The results shown here, confirm force spectroscopy studies of single aggregates as potentially very useful to characterize the thermodynamic and mechanical properties of nucleic acid-peptide complexes. On a longer term, the study of the mechanical response of aggregates related to many relevant neurodegenerative diseases could also be approached by using a peptide template instead of a DNA molecule.

CHAPTER 18

FUTURE PERSPECTIVES: EXPERIMENTS WITH AMYLOID PEPTIDES

The fact that polyanionic molecules bind and accelerate amyloid formation due to a compensation of electrostatic repulsion (see section 4.1.1.1) has been reported by several groups in the last few years [107, 129, 131, 220]. Knowledge of such widespread and nonspecific interactions is of increasing interest due to their hypothetical role on the cytotoxicity mechanism of amyloidogenic diseases (see section 4.1.1.3). As a proof of principle we have investigated the binding of the Aβ-40 peptide to dsDNA using the same DNA stretching set-up used to characterize the KF-DNA interaction (see section 14.1). The Aβ-40 peptide (4.3 kDa) is a fragmentation product of the β-amyloid protein, and is usually found in amyloid plaques of Alzheimer's disease brains. It is a less fibrillogenic variant than the presumably more cytotoxic Aβ-42 peptide (see section 4.1.1.3) [144, 271], and is therefore better suited to start exploring the interaction of amyloid peptides with DNA in force spectroscopy experiments.

We first reproduced the experimental conditions that have been previously shown to induce the association between aggregated Aβ-40 peptide and DNA in electrophoretic shift assays [272]. The peptide was dissolved in water at a 4 mM concentration and aged at 37°C to induce its aggregation [273, 274]. Immediately before performing an interaction experiment, the aged peptide was diluted to 100 μM in TBE buffer. These conditions have been reported to induce peptide-DNA complexation for Aβ-40 that has been aged in

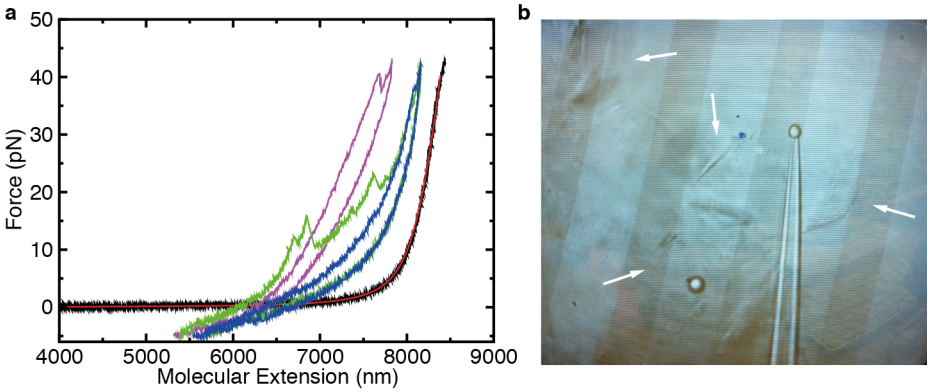

Figure 68: DNA pulling experiments with Aβ-40 peptide. (**a**) FEC of a 24-kb DNA molecule in TBE buffer before (black) and after flowing 100 μM Aβ-40 peptide for 5 min (purple, green and blue). The red line is a fit to the extensible WLC model. (**b**) Snapshot of the experimental set-up when the experiment is performed in TE 100 mM NaCl instead of TBE buffer. The higher salt concentration enhances the rate of fibril formation and experiments can not be performed due to the high abundance of large fibrils ($\geqslant$ 10 μm). This can be seen as objects indicated with arrows. In both experiments the peptide had been aged for 72 hours in water at 37°C.

water for 5 days [272]. Indeed, we expect such low salt condition[1] to (i) slow-down the self-aggregation of the Aβ-40 peptide due to electrostatic self-repulsion, and (ii) favour its electrostatic binding to the negatively charged nucleic acid.

To assess the binding of the aggregated peptide to DNA, a single half λ-DNA molecule was stretched in the optical tweezers set-up in TBE buffer and its FEC analyzed with the WLC model as in section 14.1. The freshly prepared 100 μM peptide solution was then flowed into the fluidics chamber using the microfluidics set-up with a microtube shunt (see section B.2). Although we observed the presence of aggregated species in the peptide flow after 24 hours of aging in water, we did not observe binding of Aβ-40 to DNA until we used samples that had been aged for at least 72 hours. In these experiments, the FECs obtained after flowing the amyloid peptide were very similar to those obtained for KF-DNA complexes, and showed a series of force peaks that suggest peptide-DNA complexation and compaction (Figure 68**a**). We also observed a repulsive net force at large extensions (~6 μm) probably due to the binding of long rigid amyloid fibrils that are compressed when the two beads are closer than 6 μm. On the other hand, these experiments did not show a systematic shortening of the DNA molecule during consecutive pulling cycles indicative of a transition towards an hydrophobically collapsed structure. This might be due to the higher rigidity and fibrillar structure of amyloid aggregates, or to the slower aggregation kinetics of the peptide at these specific conditions. We repeated the experiments

1 Although TBE buffer does not contain NaCl or other cations in its composition, we expect the ionic strength to be ~ 30 mM due to the partial protonation of Tris ($pK_a = 8.06$) at this buffer pH (pH ~ 8.4).

at a higher ionic strength (100 mM NaCl) to enhance its aggregation rate[2]. In these conditions it was not possible to conduct experiments with the amyloid peptide as the sample had already aggregated onto very large fibrils before flowing it into the chamber (Figure 68**b**).

As a further step, it would be interesting to perform experiments in which unaggregated peptide is flowed in a condition in which the kinetics of formation are moderately slow (hours), and wait several hours to observe if the complexation of the unaggregated peptide takes place onto the tethered DNA molecule. Another exciting but challenging experiment, would be to use recently developed polysaccharide bioorthogonal chemistry [275, 276, 277] to synthesize glycosaminoglycan molecules (e.g. heparan sulfate) modified for optical tweezers manipulation. Such templates would allow to investigate the binding of the amyloid peptides in a substrate that has been recently found to be largely associated to the cytotoxicity of amyloidogenic disorders (see section 4.1.1.3)[132]. Such a set-up would also be useful to asses the mechanism of action of certain anticancer agents that might target specific glycosylation patterns of carcinogenic cells (see section 13.2.2) [246, 245].

2 The peptide was aged for 72 hours in water at 4 mM concentration and then diluted to 100 µM in TE 100 mM NaCl instead of TBE buffer

Part IV

CONCLUSION

We shall not cease from exploration
And the end of all our exploring
Will be to arrive where we started
And know the place for the first time.

T.S. Eliot (Little Gidding)

CONCLUSION

The development of single-molecule techniques has provided scientists with a repertoire of tools that allow to measure physical quantities and scrutinize experimental systems that cannot be accessed in ensemble measurements. The ability to manipulate and observe molecular entities one at a time, has substantially modified the way in which we can approach the physics, chemistry, and biology of microscopic processes, providing an experimental playground were these traditional disciplines interlace. For instance, the stochastic nature of most experimental variables in single-molecule experiments makes useful to borrow tools from statistical physics to interpret biomolecular reactions followed with these techniques (e.g. polymer elasticity models used in chapters 7 and 14). In return, these same molecular systems under study provide amenable but yet complex experimental systems to test -or discover- new physical laws of non-equilibrium processes. Consequently, it seems likely that in the years to come, these new tools to observe and measure natural phenomena at the limit of thermal noise, will contribute to the emergence of unifying frameworks that ultimately lead to new scientific paradigms within these disciplines (physics, chemistry and biology).

A large part of this thesis is devoted to the first approach depicted earlier: using instrumentation and models based in physical principles to understand biochemical interactions. In particular, we have focused in developing tools to investigate the binding mechanisms of small ligands to DNA. We have focused in two anticancer compounds (Thiocoraline and Kahalalide F) were many bulk measurements failed to provide reproducible results likely due to the their low solubility. Understanding the mechanisms by which these molecules bind to DNA is important to determine their mechanism of action and to design new ligands with improved characteristics. Moreover, the experiments presented in these parts of the thesis (Parts ii and iii) revealed unexpected features of interest to the eye of the physicist such as metastability effects in DNA pulling experiments (chapter 7), or the existence of two different phases in a peptide-DNA compaction phenomena (chapter 14). This finding for instance has been recently observed to be particularly relevant in a very different context as it is gene delivery systems [278]. The experimental tools developed to investigate such interactions, have also been exploited to address more basic questions such as the elasticity of ssDNA and the formation of DNA secondary structures (Appendix A). Similarly, the DNA hairpin constructs and experimental protocols developed in these sections

have also been useful to test fundamental non-equilibrium theories that fall out the main scope of this thesis (see Publication list).

The first ligand investigated in this thesis (Thiocoraline) was known to bis-intercalate DNA, and to interfere with DNA replication in *in vivo* experiments. However, the main determinants of its binding to DNA (e.g. sequence preferences, kinetics) remained largely controversial. Consequently, Thiocoraline represented a good model system in which the potentialities of force-spectroscopy techniques could contribute to unravel key aspects of its binding pathway to DNA. Pulling experiments shown in chapter 7 demonstrate that the kinetics of bis-intercalation are remarkably slow, which allowed us to perform experiments in conditions of metastability. In this way we have determined the elastic properties of bis-intercalated DNA without the perturbing effects caused by force-enhanced intercalation present in previous measurements. From the elasticity measurements, we also extract equilibrium quantitates of the interaction by using classical models of nonspecific adsorption to an unidimensional lattice (MGVH model). As kinetics are an essential determinant of most therapeutic agents, we then focus in developing a methodology to study the kinetics of low solubility ligands in the optical tweezers setup (chapter 8). By combining DNA stretching and unzipping assays, we show the importance of using appropriate wash-off protocols to perform accurate kinetic measurements. Remarkably, we find that the kinetics of bis-intercalation are strongly force-dependent and experiments performed at different conditions show the existence of a mono-intercalated intermediate state. The observation of this binding intermediate also pinpoints the kind of information that many times remains hidden in bulk experiments but that can be accessed with single-molecule techniques. In this case, the ability to apply forces to the tethered DNA, allows to slow down the unbinding reaction making easier to observe the two timescales present in the experiments. The existence of such mono-intercalated intermediates had been suspected since the nearly 80's but was not observed in classical bulk assays. In this sense, the presented three-state model offers a theoretical framework from which the kinetic rates of the reaction can be extracted, and that could be useful to characterize other bis-intercalators. Similarly, we show how DNA unzipping experiments can be used to determine the binding sites of Thiocoraline, providing a single-molecule footprinting technique able to recover the preferred binding sequences of small ligands with 1 bp resolution (chapter 9). This methodology is potentially very useful as it provides direct access to the preferred binding sites due to its thermodynamic stability. Moreover, single-molecule footprinting could be applied to other complexes difficult to study in bulk such as transcription factors or multi-protein complexes binding to DNA or RNA. In this sense, chap-

ter 12 provides an overview on how this technique could be adapted to a high-throughput platform to speed up measurements.

The combination of DNA stretching and unzipping assays developed to study Thiocoraline is then used in part iii to investigate how Kahalalide F self-assembles and compacts DNA. In this sense, optical tweezers experiments with Kahalalide F represent a more complex enterprise than those with Thiocoraline, as the interplay between peptide-DNA interactions and peptide-peptide self-assembly come also into play. In this set of experiments we provide the first experimental evidence that Kahalalide F self-assembles into positively-charged nanometric particles able to bind and condense polyanions such as DNA. This interaction can be used as a model system to understand how electrostatic binding forces compete with the hydrophobic forces that drive aggregation and collapse. We show that elasticity measurements can be used to follow the initial compaction of the KF-DNA complex, and its subsequent stiffening and collapse into a highly hydrophobic structure. The existence of these two phases (electrostatic binding and hydrophobic collapse) is shown both in time-controlled and distance-controlled experiments, and the electrostatic origin of the initial binding is confirmed using uncharged peptide analogs. The combination of quantitative force-spectroscopy measurements with AFM images of the complexes and other bulk techniques (DLS, EM) provides a consistent picture of the compaction and aggregation process. In this sense, a statistical analysis similar to that previously used for DNA footprinting, can be combined with a model based on kinetic traps to determine thermodynamic parameters of the interaction such as the binding energies and barriers to disrupt the peptide-DNA condensate. Again, these results can be complemented with kinetic experiments similar to those performed for Thiocoraline, in which we follow DNA condensation at different forces. Importantly, we also show how unzipping experiments can reveal binding to ssDNA, and propose a simple methodology to study the elasticity of ssDNA using DNA hairpins. We exploit this methodology to understand how the stiffness of the polyanion affects the compaction kinetics, and later on, we also show its utility to study the elasticity of ssDNA under varying ionic conditions (Appendix A). Overall, this part of the thesis, provides a more complex framework in which the different experimental configurations, data analysis, and models developed in part ii can be used to unravel aspects of an unusual DNA-ligand interaction. This provides a first basis to investigate the self-assembly of aggregating molecules with force-spectroscopy techniques, which might be relevant to obtain quantitive data of molecular interactions related to neurodegenerative disorders (chapter 18).

Overall, in this thesis we show that single-molecule optical tweezers experiments can be used to address different types of interactions that arise between small ligands and DNA. We shows that elastic-

ity measurements using DNA pulling experiments can be used to characterize very different DNA-ligand interactions (i.e. intercalation, condensation, aggregation). Similarly, we propose different combinations of DNA stretching and unzipping assays able to tackle down the different contributions to these molecular interactions for drugs where more usual bulk techniques failed to do so. The experimental results are combined with theoretical models from which quantitative thermodynamic and kinetic parameters are obtained (e.g. free-energies of binding, kinetic rates, barriers), and that also provide a mechanistic model for the binding pathway of the ligands. We hope that the broad spectrum of results presented here might contribute to approach single-molecule force spectroscopy techniques to the chemical biology community; particularly, by developing new technologies useful for the biochemical and pharmacological industry (e.g. single-molecule footprinting) and encouraging quantitative approaches to molecular aggregation processes that are biomedically important (e.g. characterizing the interactome of amyloid-like peptides).

Part V

RESUM DE LA TESI

Quina petita pàtria
encercla el cementiri!
Aquesta mar, Sinera,
turons de pins i vinya,
pols de rials. No estimo
res més, excepte l'ombra
viatgera d'un núvol.
El lent record
dels dies
que són passats per sempre.

S. Espriu (Cementiri de Sinera)

RESUM DE LA TESI EN CATALA

Les tècniques de molècula individual permeten seguir les reaccions biomoleculars amb una resolució sense precedents, proveint als científics d'una sèrie d'instruments per a mesurar magnituds físiques i investigar sistemes experimentals difícilment accesibles amb les tradicionals mesures *en volum* (és a dir a on les mesures es realitzen amb mols de reactiu). Particularment, les pinces òptiques permeten manipular i aplicar forces a molècules individuals i determinar-ne així les seves propietats elàstiques i termodinàmiques.

L'atrapament òptic es basa en l'ús d'un feix làser focalitzat per exercir una força òptica a les microesferes (diàmetre ~ 3 μm), que queden confinades a prop del punt focal a causa de la conservació del moment lineal. Els experiments de micromanipulació es realitzen unint els extrems de la molècula que es vol estudiar a la superfícies de dues microesferes diferents, podent així aplicar forces a la molècula quan desplacem una microesfera respecte de l'altra. Per ancorar les molècules a la superfície de les microesferes s'utilitzen unions moleculars que tenen una alta afinitat (p.ex. enllaç biotina-estreptavidina). Típicament els experiments amb pinces òptiques consisteixen en la micromanipulació d'àcids nucleics (ADN, ARN) o proteïnes de forma individual. Per exemple, una molècula d'ADN pot ésser estirada per a estudiar-ne les propietats elàstiques, o oberta mecànicament (separant les dues cadenes que formen la doble hèlix) per a mesurar les energies lliures d'aparellament entre bases.

Un gran nombre d'agents anticancerígens tenen com a diana els àcids nucleics, a on s'hi uneixen a fi i efecte de dur a terme la seva acció citotòxica (p. ex. interferint amb processos cel·lulars essencials com són la replicació, la transcripció o els mecanismes de reparació). Per entendre el seu mecanisme d'acció és important conéixer en quines posicions, amb quina afinitat, i amb quina cinètica s'uneixen a diferents seqüències d'ADN. Els experiments de molècula individual amb pinces òptiques permeten determinar la termodinàmica i cinètica d'unió de molts d'aquests lligands, especialment aquells difícils de caracteritzar mitjançant mesures *en volum*. És per això que un dels objectius principals d'aquesta tesi, ha estat aprofitar les potencialitats de les mesures de molècula individual per a caracteritzar pèptids anticancerígens poc solubles i difícils d'estudiar amb tècniques alternatives: des de la cinètica i termodinàmica d'unió, a l'especificitat en seqüència i la cinètica d'autoacoblament.

La Tiocoralina és un d'aquests pèptids, i s'uneix a l'ADN a través d'un procés conegut com a bis-intercalació. Els intercalants són petits

compostos planars que s'uneixen a l'ADN insertant-se entre parells de bases adjacents. En un bis-intercalant dos d'aquests anells intercalants es troben unitats per una cadena flexible, i la molècula per tant s'insereix entre dos parells de bases consecutius (de forma similar a una grapa), desenrotllant i allargant l'ADN una distància ~6.8 Å. Tot i que es sap que la Tiocoralina interfereix amb la replicació de l'ADN *in vivo*, el seu mecanisme específic d'unió continua sent poc conegut.

Els experiments amb pinces òptiques mostren que la cinètica d'intercalació és molt lenta (hores) i que té una forta dependència amb la força: mentre que l'aplicació de força accelera la unió del lligand, també n'alenteix la cinètica de sortida. Combinant els experiments d'estirament molecular amb els d'obertura mecànica de l'ADN mostrem la importància d'utilitzar protocols de neteja apropiats per a dur a terme aquestes mesures cinètiques de forma acurada. Així mateix, experiments cinètics duts a terme en diferents condicions, mostren que la unió de la Tiocoralina succeeix a través d'un estat intermedi mono-intercalat que causa la cinètica extremadament lenta del lligand. La observació d'aquest intermedi remarca el tipus d'informació que molts cops roman inaccesible en la majoria de mesures *en volum*, però que pot ser observat amb tècniques de molècula individual. A través de la observació d'aquest intermedi, es desenvolupa un model cinètic de tres estats que permet explicar totes les observacions experimentals i que pot ser d'utilitat per a caracterizar altres bis-intercalants. La cinètica extremadament lenta de la Tiocoralina, també permet caracteritzar les propietats elàstiques de l'ADN bis-intercalat. Mesures prèvies d'aquests paràmetres es trobaven afectades pel fet que el nombre d'intercalants units depèn fortament en la força aplicada, conduint a conclusions errònies. En el cas de la Tiocoralina, es poden realitzar experiments d'estirament molecular fins a forces altes en condicions de metastabilitat, i obtenir-ne així les seves propietats elàstiques de forma acurada. D'aquestes mesures, també n'extreiem paràmetres d'equilibri de la interacció utilitzant models d'absorció no-específica en una xarxa unidimensional (model de McGhee i Von Hippel). Així mateix, també mostrem que els experiments d'obertura mecànica de l'ADN es poden utilitzar per a determinar les seqüències d'unió preferides de la Tiocoralina anb una resolució d'un parell de bases, obtenint així "l'empremta" molecular del lligand a través d'una tècnica de molècula individual. Aquesta metodologia té la singularitat de mostrar les seqüències preferides del lligand en base a la seva estabilitat termodinàmica i sense requerir l'ús d'enzims o mètodes radioactius. Així mateix mostrem que la tècnica es pot paral·lelitzar mitjançant instrumentació de pinces magnètiques, i que per tant podria tenir potencials aplicacions per estudiar factors de transcripció o complexos multi-moleculars difícils de caracteritzar amb mètodes convencionals.

Aquesta combinació d'experiments d'estirament molecular i d'obertura mecànica de l'ADN també es pot utilitzar per seguir el procés d'autoacoblament del pèptid Kahalalide F, i la compactació que aquest indueix a l'ADN. En aquest cas, els experiments amb pinces òptiques amb el Kahalalide F representen una empresa més complexa que amb la Tiocoralina, ja que hi ha un equilibri entre les interaccions d'unió pèptid-ADN i les interaccions d'agregació i autoacoblament pèptid-pèptid. En aquesta sèrie d'experiments obtenim la primera demostració experimental que el Kahalalide F forma partícules nanomètriques carregades positivament que poden unir-se i condensar polianions com l'ADN. Aquesta interacció per tant pot ser utilitzada com un sistema model per entendre com les forces electrostàtiques d'unió competeixen amb les forces hidrofòbiques que porten a l'agregació i el col·lapse. Les mesures d'elasticitat mostren que la interacció té dues fases: una compactació inicial d'origen electrostàtic, i una subsegüent rigidificació i col·lapse del complex en una estructura altament hidrofòbica. L'existència d'aquestes dues fases (unió electrostàtica i col·lapse hidrofòbic) es pot observar tant en experiments en què s'utilitza la distància o el temps com a paràmetre de control. Així mateix, experiments realitzats a diferents concentracions de sal, i amb anàlegs del pèptid no-carregat, mostren l'origen electrostàtic de la interacció. La combinació de mesures quantitatives d'espectroscòpia de forces amb imatges de microscopia de força atòmica i tècniques complementàries, proporcionen un marc consistent del procés d'agregació i compactació. La modelització dels resultats experimentals mitjançant un model de trampes cinètiques permet determinar paràmetres termodinàmics de la interacció com són l'energia d'unió i la barrera cinètica per a separar el complex. Aquests resultats es complementen amb experiments cinètics per seguir la condensació de l'ADN a diferents forces, i amb experiments d'obertura mecànica de l'ADN que mostren que el Kahalalide F també interactua amb la cadena simple d'ADN. A fi d'entendre aquesta interacció en detall, proposem una tècnica senzilla per a estudiar l'elasticitat de la cadena simple d'ADN mitjançant els experiments d'obertura mecànica. Primerament utilitzem aquesta metodologia per a estudiar l'efecte de la rigidesa del polianió en la cinètica de compactació, i posteriorment demostrem que aquest mètode permet estudiar les propietats elàstiques de la cadena simple d'ADN en diferents condicions iòniques. En conjunt, aquests resultats constitueixen una primera base per a estudiar els processos d'agregació i autoacoblament molecular mitjançant tècniques d'espectroscòpia de forces. Aquestes tècniques poden ser particularment rellevants per a obtenir informació quantitativa d'interaccions moleculars relacionades amb malalties neurodegeneratives. En aquest sentit, la utilitat de la metodologia desenvolupada, s'explora amb pèptids amiloides relacionats amb la malaltia d'Alzheimer.

Part VI

APPENDIX

CHAPTER A

ELASTIC PROPERTIES OF dsDNA AND ssDNA MOLECULES

A.1 ELASTIC MODELS OF LINEAR POLYMERS

A.1.1 *Freely Jointed Chain (FJC) model*

The simplest model to treat the elasticity of a long, linear polymer is the Freely Jointed Chain (FJC) model. In this model the molecule is considered as a chain of perfectly stiff subunits that are linked together by totally free hinges (allowing free rotations in all directions) (Figure 69**a**). In the FJC model, the energy E_{FJC} of the chain in the presence of an external force $\vec{F}$ in the $\hat{z}$ direction is given by:

$$E_{FJC}(\{\hat{t}_i\}) = -\sum_{i}^{N} Fb\hat{t}_i \cdot \hat{z} = -Fb\sum_{i}^{N} \cos\theta_i \tag{48}$$

where each of the N monomers has a length b and points in the direction of $\hat{t}_i$.

The molecule explores its configurational space due to thermal fluctuations, and in the absence of external forces it performs a random walk as all configurations have an equal energy. However, when an external force is applied to its ends, there is a free energy cost associated with forcing the chain to adopt an entropically less likely

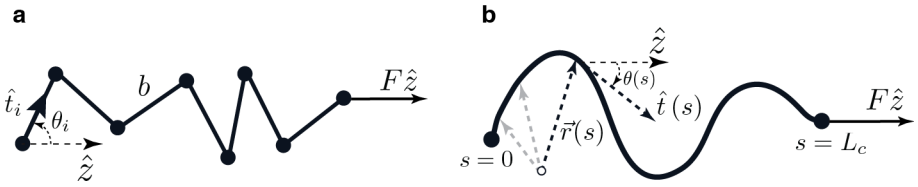

Figure 69: Sketch of the FJC and WLC models. (a) The FJC model represents the polymer as a collection of N identical uncorrelated segments of length b, each pointing in the direction of the vector $\hat{t}_i$ and forming and angle θ_i with the direction of the stretching force $\hat{z}$. **(b)** The WLC model represent the polymer as a continuous elastic chain that has a bending rigidity, and whose configuration can be described in terms of the position vector $\vec{r}(s)$ along its contour length s.

conformation. The partition function of the system is then obtained by integrating over all possible configurations:

$$Z_N = \prod_i^N \int_0^{2\pi} d\phi_i \int_0^{\pi} e^{Fb\cos\theta_i/k_BT} \sin\theta_i \, d\theta_i = 4\pi \frac{k_BT}{Fb} \sinh\left(\frac{Fb}{k_BT}\right) \tag{49}$$

From this result, the average end-to-end distance $\langle z \rangle$ along the stretching direction can be directly found:

$$\langle z \rangle = k_BT \frac{\partial \ln Z_N}{\partial F} = Nb\left[\coth\left(\frac{Fb}{k_BT}\right) - \frac{k_BT}{Fb}\right] \tag{50}$$

This force-extension relation for the entropic elasticity of a polymer that follows the FJC model is usually expressed as:

$$x(F) = L_c\left[\coth\left(\frac{Fl_k}{k_BT}\right) - \frac{k_BT}{Fl_k}\right] \tag{51}$$

where x is the end-to-end distance in the direction of the applied force F, L_c is the molecular contour length ($L_c = Nb$), and l_k is the Kuhn length ($l_k = b$) of the characteristic monomer.

As the monomers are inextensible, at a high stretching force the polymer become fully aligned and the molecular extension x saturates at the contour length value L_c. However, real polymers have a certain deformability and can be extended beyond their contour length. The FJC model can be modified to account for this extensibility by introducing a stretch modulus S:

$$x(f) = L_c\left[\coth\left(\frac{Fl_k}{k_BT}\right) - \frac{k_BT}{Fl_k}\right]\left(1 + \frac{F}{S}\right) \tag{52}$$

The FJC model is well-suited to characterise flexible polymers that have a low bending energy ($E_{bend} << k_BT$) such as ssDNA. In this case, a fit of an experimental FEC to the model yields values of l_k that are similar to the the characteristic monomer length (e.g. the interphosphate distance for ssDNA), reflecting that the model gives a

good physical description of the polymer. However, for stiffer polymers such as dsDNA, where base-stacking interactions and hydrogen bonds stabilise the double helix conformation, this model does not fit well the experimental FECs even if a larger effective l_k is used. For such molecules, a model that accounts for the bending rigidity of the polymer must be used.

A.1.2 *Worm Like Chain (WLC) model*

The WLC is a semi-flexible polymer model that describes the molecule as a continuous inextensible rod (Figure 69**b**) that requires a certain energy to be bent. Therefore the model considers that the energy of the stretched polymer is given by an integral of two terms: the resistance of the chain to bending, and the previous entropic term due to the application of a force F that aligns the molecule [279]:

$$E_{WLC} = \int_0^{L_c} \left(\frac{A}{2} \left| \frac{d\hat{t}(s)}{ds} \right|^2 - F\cos\theta(s) \right) ds \tag{53}$$

where $\hat{t}(s)$ is the unit tangent vector to the chain and $\cos(s)$ is the angle between $\hat{t}(s)$ and $\hat{z}$. In this model, the persistence length l_p is the parameter that controls the elasticity of the polymer and is related to the factor A of the bending energy through: $l_p = A/k_bT$. Consequently, l_p gives a measure of how susceptible is the molecule to bending by thermal fluctuations, and is the characteristic length defining the correlation distance of the polymer at zero force ($< \hat{t}(0) \cdot \hat{t}(s) > \sim e^{-|s|/l_p}$).

The partition function Z can then be calculated to derive a force-extension relation analogously as done for the FJC model. However, the resulting expression can not be solved analytically, and data is usually fitted by means of the interpolation formula proposed by Marko and Siggia [280, 281]:

$$F(x) = \frac{k_bT}{l_p} \left[\frac{1}{4(1-x/L_c)^2} - \frac{1}{4} + \frac{x}{L_c} \right] \tag{54}$$

This model fits well dsDNA in the low force regime ($F > 10pN$) but fails at higher forces due to the deformability of the DNA double helix, that can be extended over its contour length. As done for the FJC model, we can introduce a stretch modulus S that accounts for the extensibility of the monomers. This is done by replacing the total contour length L_c by a term $L_c(1 + F/S)$. This enthalpic correction gives the extensible WLC model that describes well the experimental FECs of dsDNA up to forces of $< 50pN$:

$$F(x) = \frac{k_bT}{l_p} \left[\frac{1}{4(1-x/L_c+F/S)^2} - \frac{1}{4} + \frac{x}{L_c} - \frac{F}{S} \right] \tag{55}$$

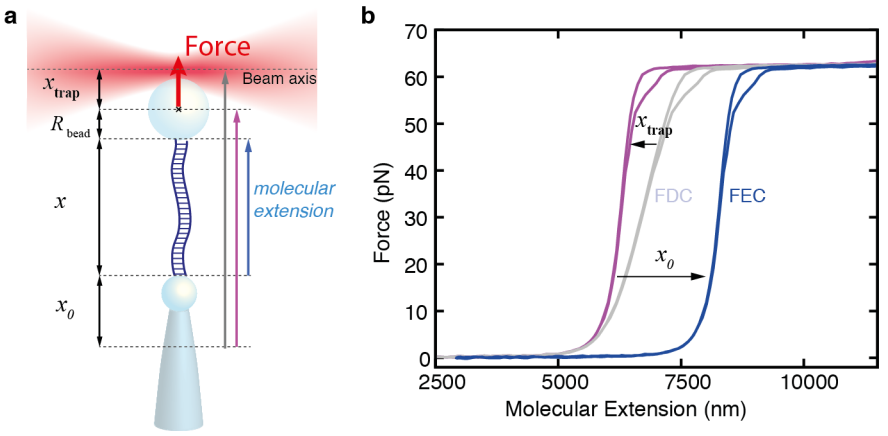

Figure 70: From FDCs to FECs. (a) Schematics of the set-up. The total measured distance (position of the optical trap) is expressed as the sum of the molecular extension (x), the trap elastic contribution (x_{trap}), and a constant offset (x_0) that includes all other contributions. **(b)** In a typical DNA pulling/unzipping experiment the FDC is measured (gray). Conversion from distance to molecular extension is done by subtracting first the compliance of the optical trap x_{trap} (purple), and then the distance offset x_0 (blue).

Finally, both the inextensible and extensible interpolation formulas of the WLC model can be improved by adding a polynomial correction $\mathcal{O}(x/L_c)^7$ that reduces the error below 0.01% [248]. This is the expression of the WLC formula used throughout this thesis that we synthesize below:

$$F(l) = \frac{k_B T}{L_p}\left(\frac{1}{4(1-l)^2} - \frac{1}{4} + l + \sum_{i=2}^{i\leqslant 7} \alpha_i l^i\right) \tag{56}$$
$$l = x/l_0 - F/S(\text{extensible})$$
$$l = x/l_0(\text{inextensible})$$

A.2 FROM EXPERIMENTAL FDCs TO FECs

In the minitweezers set-up we obtain a direct measurement of the force applied to the tethered molecule (F) and of the relative position of the optical trap (x_{trap}). Relative displacements of the position of the optical trap are directly measured from the deflection of the laser beam by means of position sensing detectors (PSD's) with an accuracy of 1 nm. To infer the absolute molecular extension of the tethered molecule (end-to-end distance x) the measured distance must be corrected by two factors. First the elastic contribution of the optical trap must be subtracted to obtain the relative molecular extension; then the absolute molecular extension is inferred by subtracting a distance offset x_0 (Figure 70**a**).

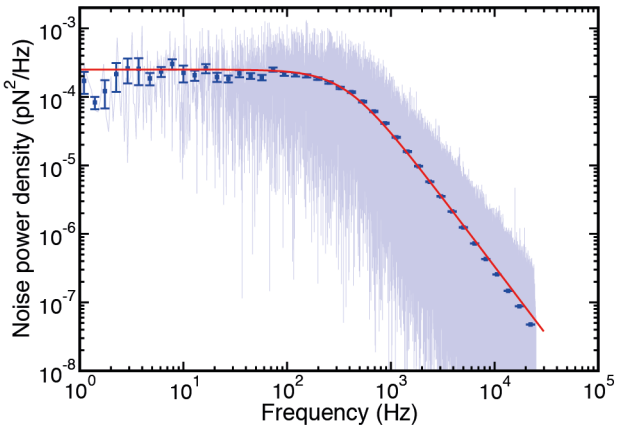

Figure 71: Trap stiffness determination. Power spectral density distribution of force fluctuations using a bead sized 3 μm in diameter. Ten seconds output was recorded at 50 kHz. An exponential average is shown in dark blue, and a fit to Eq. 57 in red.

A.2.1 *Trap stiffness correction and zero-force baseline*

In the explored range of forces, the optical trap is well described by an harmonic well ($V \sim \frac{1}{2}kx_{trap}^2$). Consequently, the optical trap compliance can be corrected by subtracting the quantity F/k (where k is the trap stiffness) from the distance measurement (Figure 70**b**).

The trap stiffness can be determined by recording the Brownian motion of the trapped bead at 50 kHz with a high-frequency data acquisition system directly connected to the electronic board of the optical tweezers as described in [257]. The power spectrum of the force fluctuations is calculated using a fast fourier transform (FFT) algorithm (Figure 71), and fitted to the following Lorentzian:

$$< \Delta F^2(\omega) >= \frac{2k_B T k^2}{\gamma(w_c^2 + w^2)} \tag{57}$$

where k is the trap stiffness, γ is the drag coefficient ($\gamma = 6\pi\eta R$ for a spherical particle), and w_c is the corner frequency ($w_c = k/\gamma$). The fits are performed using a Levenberg-Marquadt algorithm with free-parameters k and γ. Using this method the trap stiffness was found to be 0.07 ± 0.005 pN/nm (N=6 beads).

The force measurement is not completely homogenous when the optical trap is moved over a long distance range using the "wigglers" (i.e. piezoelectric actuators that bend the tip of the optical fiber). Consequently, a zero-force baseline that can be subtracted to the data is usually recorded by moving a trapped bead along the full distance range (~10 μm) in the pulling direction (Figure 72). Force deviations from the zero value are small (typically below 2 pN) and are only important for pulling experiments involving long DNA constructs (a few μm long). In the experiments shown in this thesis, this correction is only relevant to perform accurate elasticity measurements of

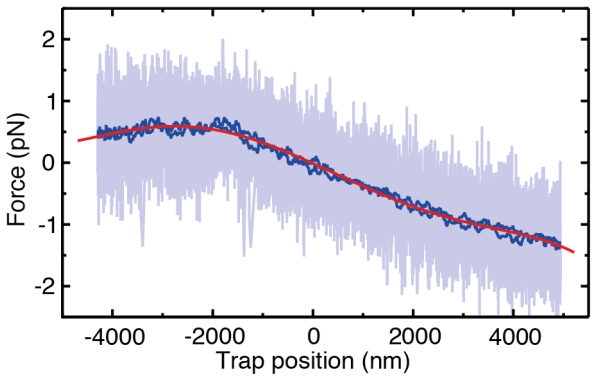

Figure 72: Zero-force baseline. A trapped bead is moved along the full range of the piezoelectric actuators to measure the zero-force baseline. Data is recorded at 1 kHz (light blue), filtered to 10 Hz (dark blue) and fitted to a polynomial function (red). Pulling speed 200 nm/s

dsDNA and ssDNA. Particularly, the high non-linearity of the elastic response of dsDNA (see Figure 70**b**) makes very important to subtract this zero-force correction to the FECs. In fact, to fit the FECs of dsDNA to the WLC model, we also added an extra fitting parameter f_0 that could take values between ± 0.5 pN. This fitting parameter allows to fine tune the total shift of this zero-force correction, and provides more consistent fits to the WLC model.

A.2.2 *Relative molecular extension correction for ssDNA*

To determine the distance offset in the unzipping and ssDNA experiments, the zero extension is found by bringing the two beads in close contact. This corresponds to $x = 0$ nm, and the DNA hairpin should be fully formed at this extension. Then the DNA hairpin is unzipped at a slow pulling speed ($\leqslant$ 100 nm/s). The DNA unzipping pattern is a fingerprint of the DNA sequence, and therefore we can determine if the DNA molecule was tethered in an optimal configuration by checking that the full unzipping pattern of the molecule is seen. The distance offset can then be corrected, or the molecule directly discarded if the initial region of the hairpin is not visible due to non-optimal tethering. This allows us to determine the absolute molecular extension with an accuracy below 5%. For experiments with the long DNA hairpin (6.8 kb), the contribution of the DNA handles and the blocking loop oligo to the total molecular extension is omitted as it corresponds to approximately 20 nm (well below 1% of the total molecular extension).

A.2.3 *Relative molecular extension correction for dsDNA*

Similarly, for the DNA stretching experiments, the absolute molecular extension can be found by bringing the two beads in close contact. However in this set-up it is very likely to obtain multiple tethers while the two beads are close to each other. For these experiments a video-based method is preferred. The absolute molecular extension is obtained by measuring the bead-to-bead distance from the CCD camera images. The measurement is usually performed at a constant force of 5-10 pN, and the value determined in this way is used to correct the relative measurement distance obtained from the light-lever measurement. Indeed, at some point it would be interesting to implement a bead-tracking system to automatize this alternative method to measure molecular extensions.

Still, for the DNA stretching experiments included in this thesis it was not always needed to perform such offset measurements. The half λ-DNA molecules used for these experiments were synthesized by attaching to each end of the DNA molecule short oligonucleotides labelled with digoxigenins or biotins. This method allows to know with higher accuracy the expected contour length of the tethered DNA molecule, at a price of a lower linkage strength than by using PCR-based methods. By assuming a 0.32 nm/bp distance, the contour length of the DNA molecule l_0 can be inferred, and the offset x_0 obtained from a fit to the WLC model (Eq. 55[1]) using as free parameters l_p, x_0 and S.

A.3 ELASTIC PROPERTIES OF ssDNA AT MONOVALENT AND DIVALENT SALT CONDITIONS

Understanding the elastic behavior of single-stranded DNA (ssDNA) is of great relevance because of its major role in many biological processes, such as replication, recombination, repair, transcription and transposition of DNA [282]. In particular characterizing its elastic response and the formation of secondary structure is an important task in comprehending its physicochemical properties. Despite the large number of papers in the single molecule field, the existing literature addressing the mechanical properties of ssDNA is scarce especially when compared to that focused on dsDNA [283]. Probably this is due to the fact that ssDNA is difficult to manipulate because it folds into non-specific secondary structures [284], and binds to surfaces (e.g. glass, polystyrene) due to the hydrophobicity of the exposed bases [285, 286, 287, 288]. Using the blocking oligo approach explained in section 15.2.1 we have studied the elastic properties of ss-

1 were x is replaced by $x - x_0$

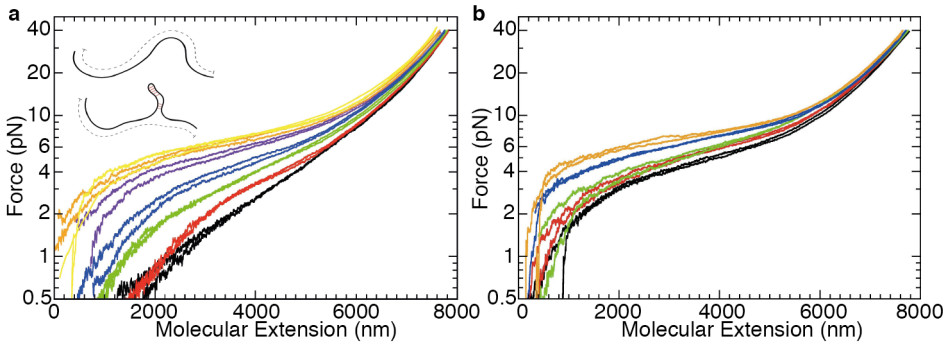

Figure 73: FEC of 13-kb ssDNA at varying salt conditions. (**a**) FEC at varying NaCl conditions: 10 mM (black), 25 mM (red), 50 mM (green), 100 mM (blue), 250 mM (purple), 500 mM (orange), 1000 mM (yellow). In the inset, pictorial representation showing how the formation of secondary structure reduces the apparent contour length of the molecule (dashed lines). (**b**) FEC at varying $MgCl_2$ conditions: 0.5 mM (black), 1 mM (red), 2 mM (green), 4 mM (blue), 10 mM (orange). Both figures show two release cycles at each condition to illustrate the variability between experiments. Data is obtained at a pulling speed of ~30 nm/s and filtered at 2 Hz bandwidth.

DNA under varying monovalent and divalent salt conditions[2] [197]. This knowledge contributes to a better understanding of the elastic parameters describing ssDNA and should be most useful to extract the base pairing free energies of dsDNA in the presence of divalent ions from mechanical unzipping experiments.

The FEC of a 13-kb ssDNA molecule in the presence of a varying concentration of monovalent ions (NaCl) is show in Figure 73**a**. At low forces (< 10 pN) and high salt (> 100 mM) a plateau compatible with the formation of secondary structure is observed [289, 290]. The formation of secondary structure at low forces reduces the effective length of the tethered molecule (Figure 73**a**, pictorial draw), leading to the observed force plateau. The increase in the height of the plateau at higher salt concentration shows that base pairing is stabilized with salt [291]. Moreover, at high salt concentration the Debye screening length becomes shorter ($\simeq 0.3$ nm at 1 M NaCl), and DNA-backbone electrostatic self-repulsion cannot prevent non-native base pair formation. In experiments performed with a varying concentration of divalent ions (Figure 73**b**) a similar trend is observed, although the relative height of the plateau is larger than for monovalent ions. For instance, at 10 mM $MgCl_2$ the height of the plateau is comparable to

2 Experiments with monovalent ions were performed at room temperature (25°C) in a buffer containing TE pH7.5 (Tris.HCl 10 mM, EDTA 1 mM), 0.01% NaN_3 and a varying concentration of NaCl (10, 25, 50, 100, 250, 500 and 1000 mM). Experiments with divalent ions were performed at 25°C in a buffer containing 10mM Tris.HCl, pH 7.5, 0.01% NaN_3 and a varying concentration of $MgCl_2$ (0.5, 1, 2, 4, 10 mM). An additional amount of 8 mM of monovalent cations contributed by the buffer (Tris protonation and NaN_3) must be added to all NaCl concentrations and considered as potentially competing with divalent cations in the $MgCl_2$ case. However, as discussed below, the latter effect turns out to be negligible in the range of divalent salt concentrations studied in this paper.

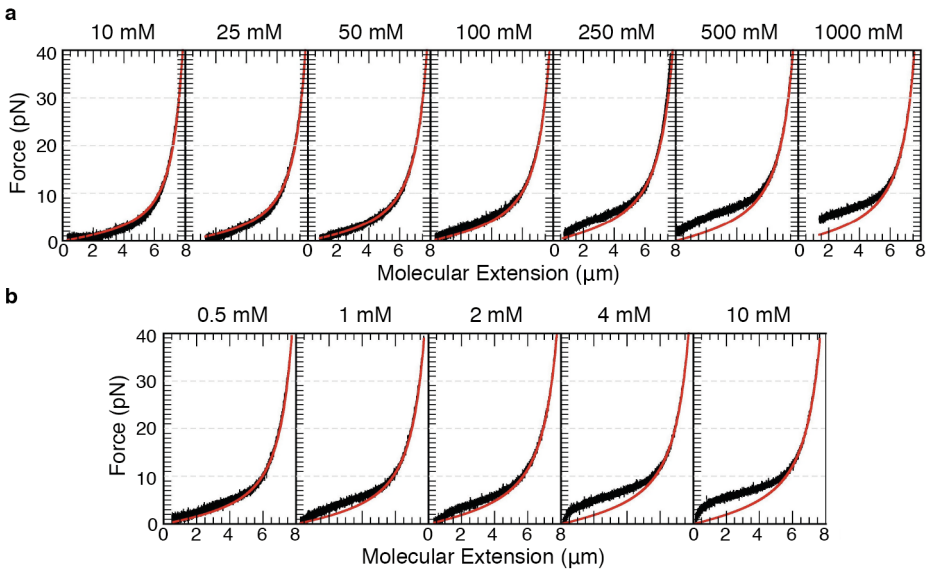

Figure 74: Elastic response and fit to the WLC model at various salt conditions. (**a**) FEC at varying NaCl concentrations (black) and fit to the inextensible WLC model (red). (**b**) FEC at varying $MgCl_2$ concentrations (black) and fit to the inextensible WLC model (red). Data is obtained at a pulling speed of ~30 nm/s and filtered at 400 Hz bandwidth. Fits are performed on the force range 10-40 pN.

that observed for 1 M NaCl. Moreover, for $MgCl_2$ the plateau extends down to lower extensions ($\leqslant$ 4 μm) falling to zero force more steeply at very low extensions. This effect is neat at high divalent salt conditions (Figure 73**b**); in contrast, in the other extreme conditions, the equivalent FECs at low monovalent salt for NaCl (e.g. 10 mM) hardly exhibit a plateau (Figure 73**a**).

To quantify the elastic properties of ssDNA, we fitted the FECs to the inextensible WLC model (Eq. 54) using the persistence length l_p and contour length L_c of the molecule as fitting parameters (Figure 74 and Table 3). Even if our molecular construct contains a short stretch of dsDNA (88-bp) serially linked to the ssDNA, the contribution of these dsDNA parts to the measured elastic response of the ssDNA is expected to be negligible (less than 1%) due to its much higher stiffness (50-fold per base pair) and much shorter length. FECs have been fitted in the range 10-40 pN in order to avoid the presence of secondary structure. In fact secondary structure formation should be suppressed at forces above 10-15 pN, large enough to unzip DNA. The WLC model describes well the curves in the range of the fit, but fails to describe the FEC at forces lower than 10 pN. In particular at these forces, the formation of secondary structure induces positive force deviations from the WLC at concentrations that are $>$ 100 mM for monovalent salt and in the whole range of divalent salt concentrations. Moreover, at monovalent salt concentrations that are $<$ 50 mM, some negative deviations at forces $<$ 10 pN are also observed. These can be attributed to excluded volume effects [290, 292, 293].

[NaCl] (mM)	l_p (nm)	St.Dev. (nm)	L_c (nm/base)	St.Err. (nm/base)
10	1.19	0.06	0.68	0.02
25	1.04	0.06	0.69	0.02
50	0.99	0.05	0.68	0.02
100	0.87	0.04	0.69	0.03
250	0.81	0.04	0.69	0.03
500	0.78	0.04	0.69	0.02
1000	0.76	0.05	0.70	0.02
$[MgCl_2]$ (mM)	l_p (nm)	St.Dev. (nm)	L_c (nm/base)	St.Err. (nm/base)
0.5	0.93	0.06	0.69	0.03
1	0.86	0.04	0.69	0.02
2	0.78	0.04	0.70	0.02
4	0.79	0.05	0.70	0.02
10	0.75	0.04	0.70	0.02

Table 3: Elastic parameters of ssDNA obtained from a fit of the FECs to the WLC model (l_p and l_0) at varying NaCl and $MgCl_2$ concentrations. A fixed amount of 8 mM of monovalent salt contributed by the buffer must be added to each condition.

In fact, at low concentrations, the Debye length can be larger than the persistence length of ssDNA so electrostatic repulsion between different DNA segments inhibits the bending of the filament. Finally, in the explored range of forces, using the extensible WLC (Eq. 55) does not appreciably improve the fits and the discrepancy on the parameters is below 3% for l_p and 1% for L_c. The FECs have also been fitted to the extensible FJC model (Eq. 52) in the same range of forces (Figure 75 and Table 4). Similarly to previous studies [249], fits of the FECs were performed by fixing the contour length to the crystallographic length l of ssDNA ($l = 0.57$ nm or L_c=7857 nm for a 13650 bases molecule) [290], and using the Kuhn length l_k and stretch modulus S as fitting parameters. Our values for the Kuhn length (Table 4) agree with those previously reported, which fall in the range 1.5-1.6 nm at [NaCl]=150 mM and in phosphate buffered saline (PBS) [249, 294, 295, 296]. The stretch modulus shows a non-linear trend with concentration and ranges between 650-850 pN, being also compatible with previous results [249, 294, 295]. Moreover, the values that we found for the Kuhn length and persistence length using both models follow the relation $l_K = 2l_p$, as expected for semi-flexible polymer models when the value of l_p is comparable to the segment length l.

It is known that the persistence length l_p of a semi-flexible polymer can be expressed as the sum of two terms:

$$l_p = l_p^0 + l_p^{el} \tag{58}$$

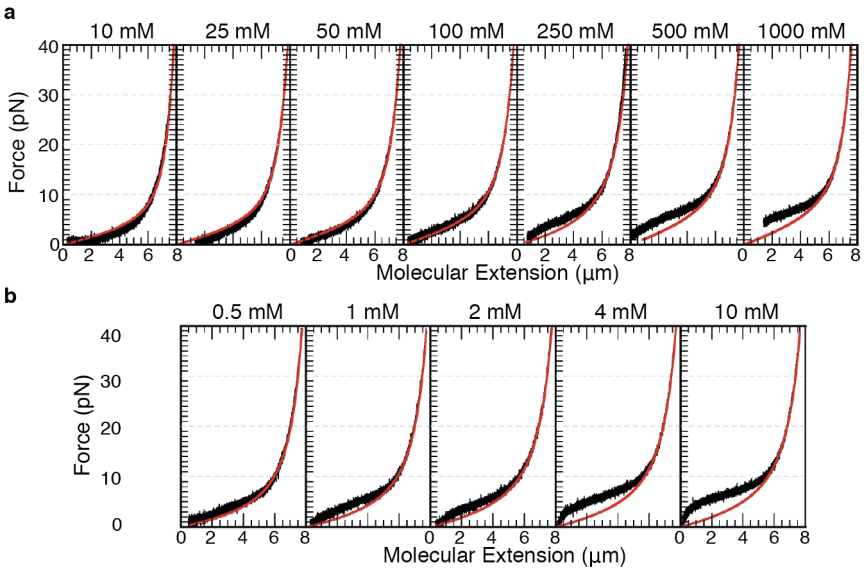

Figure 75: Elastic response and fit to the FJC model at various salt conditions. (**a**) FEC at varying NaCl concentrations (black) and fit to the extensible FJC model (red). (**b**) FEC at varying $MgCl_2$ concentrations (black) and fit to the extensible FJC model (red). Data is obtained at a pulling speed of ~30 nm/s and filtered at 400 Hz bandwidth. Fits are performed on the force range 10-40 pN.

where l_p^0 is the intrinsic persistence length and l_p^{el} is an electrostatic contribution that depends on the Debye screening length and consequently on the inverse of the square root concentration of ions in solution ($l_p^{el} \propto \frac{1}{[Ion]^{\nu/2}}$) [297, 298, 299, 300, 301]. Figure 76**a** shows l_p as a function of the inverse of the squared root of salt concentration. Remarkably, we observe a linear dependence for both types of salts that collapse into a single linear trend if a factor of 100 between both ions is considered ($FEC_{[MgCl_2]}$ ~$FEC_{100[NaCl]}$).

These results confirm the hypothesis that magnesium is roughly 100 times more efficient than sodium in non-specifically screening DNA and affecting the elastic properties of ssDNA. As has been previously suggested [302], such factor between monovalent and divalent salt concentrations cannot be explained by a Debye-Hückel picture of a diffuse cloud of screening counterions. However using strongly correlated liquid models of counterion/macroions interaction [303] a factor ranging from 50 to 200 is expected [302]. If we perform a fit with ν as a fitting parameter we then obtain ν=1.15±0.25 in the case of NaCl and ν=1.65±0.85 in the case of $MgCl_2$. Our results are in agreement with previous force spectroscopy measurements obtained by chemical denaturation of a 3-kb dsDNA molecule (Figure 76**a**, blue circles), albeit the present method provides significantly lower standard errors. These findings are also consistent with those reported in the folding of a ribozyme [304], and in pulling experiments of RNA hairpins [305, 306]. Figure 76**a** also shows the results of a fit with ν=1. For both salts, a value of l_p^0 equal to 0.7 nm is obtained. This value

[NaCl] (mM)	L_K (nm)	St.Dev. (nm)	S (pN)	St.Dev. (pN)
10	1.91	0.04	760	50
25	1.78	0.06	670	50
50	1.68	0.05	700	50
100	1.55	0.04	630	70
250	1.43	0.06	707	60
500	1.38	0.05	710	50
1000	1.31	0.06	850	100
$[MgCl_2]$ (mM)	L_K (nm)	St.Dev. (nm)	S (pN)	St.Dev. (pN)
0.5	1.59	0.06	710	60
1	1.50	0.03	700	50
2	1.41	0.03	630	60
4	1.41	0.05	670	50
10	1.34	0.04	710	60

Table 4: Elastic parameters of ssDNA obtained from a fit of the FECs to the extensible FJC model (l_k, S) at varying NaCl and $MgCl_2$ concentrations. A fixed amount of 8 mM of monovalent salt contributed by the buffer must be added to each condition.

is in good agreement with the previously reported values for single-stranded nucleic acids, in particular for poly-U (0.67 nm) [307, 301] and for chemically denatured ssDNA (0.6 nm) [293, 292].

Interestingly enough these results show that divalent cations have a notably greater effect in screening electrostatic charge than monovalent ions, as has been observed in previous thermal- and tension-induced base-pair melting experiments. Indeed, the intrinsic persistence length is ~ 0.7 nm for both types of salt and the effectivity of divalent cations in screening electrostatic interactions appears to be 100-fold as compared to monovalent salt, in line with what has been recently reported for single-stranded RNA [305]. The values of the Kuhn length obtained from the extensible FJC model display a similar behaviour (Figure 76**b**), in agreement with these findings. The contour length of the ssDNA shows a weak dependence with salt concentration (Table 4), suggesting that the variation of the ionic force of the medium does not affect the conformation of the sugars in the force range explored; in fact, a fraction of the deoxyriboses could interconvert from C3′-endo (interphospate distance 0.6 nm) to C2′-endo conformation (interphospate distance 0.7 nm) resulting in a change in the contour length [308, 309]. The WLC is sufficient to describe phenomenologically the elastic behaviour of ssDNA at forces below 40 pN but only if the contour length of the ssDNA is a fitting parameter to the elastic model. In contrast, if we fix the contour length of the ssDNA using the crystallographic interphosphate distance, then

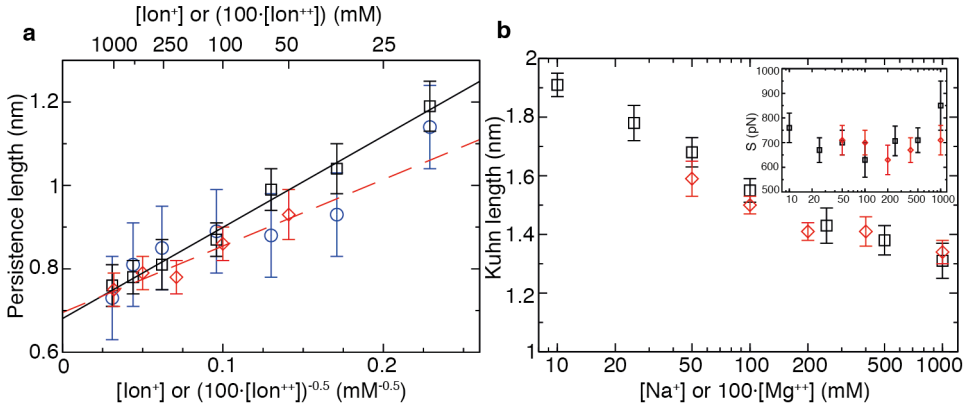

Figure 76: Salt dependence of ssDNA elastic parameters. (**a**) Persistence length values for ssDNA at different salt conditions in NaCl (black squares) and $MgCl_2$ (red diamonds). In blue circles we show data obtained from [62]. Magnesium concentrations along the x-axis have been multiplied by a factor 100. Solid black and dashed red lines correspond to the linear fits to NaCl and $MgCl_2$ data. The parameters obtained from the fit $l_p = l_p^0 + A \cdot [Ion]^{-0.5}$ are: $l_p^0 = 0.68 \pm 0.04$ and $A = 2.19 \pm 0.10$ ($\chi^2 = 0.67$) in the case of NaCl; and $l_p^0 = 0.70 \pm 0.04$ and $A = 1.6 \pm 0.6$ ($\chi^2 = 0.70$) in the case of $MgCl_2$. (**b**) Kuhn length and stretching modulus (Inset) of ssDNA versus salt concentration for NaCl (black squares) and $MgCl_2$ (red diamonds) in normal-log scale. Magnesium concentrations have been multiplied by a factor 100.

we must use an extensible model (such as the extensible FJC) to accurately fit the FECs.

A.3.1 *Secondary structure formation in ssDNA*

The blocking oligonucleotide used to inhibit the formation of the hairpin (see section 15.2.1) does not prevent the formation of non-native secondary structure at low enough forces (which are driven by stacking and base pairing interactions). In fact, the shoulder observed in the FECs at high salt concentrations suggests that ssDNA collapses in a condensed phase. This effect is more prominent in the presence of divalent cations indicating the successful coordinated action of the doubly charged cations to bring distal segments of ssDNA close to each other.

Despite the fact that we do not have a phenomenological model describing this condensed phase, the large number of structural motifs that can be formed and their low energetic stability produces a soft shoulder in the FEC without any visible force rips. The competition between the large number of possible configurations that can be formed in such condensed phase suggests that configurational entropy (rather than enthalpy) is the main driving force toward the formation of secondary structure that prevents the reformation of the native hairpin. Several effects contribute to the soft shoulder observed in the FEC: the large flexibility of the ssDNA due to its low persistence

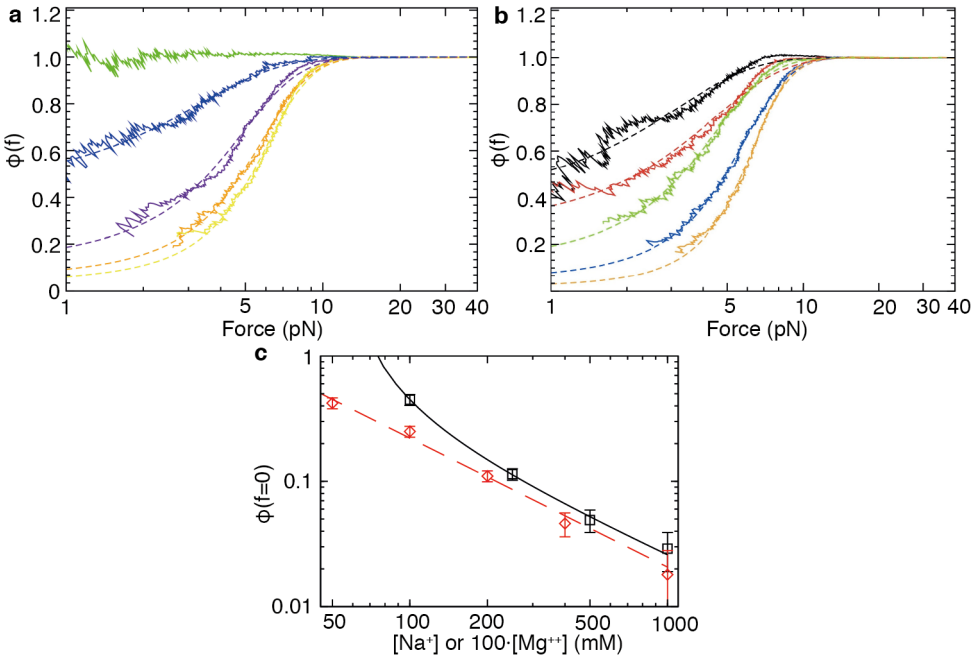

Figure 77: Secondary structure formation. (**a**) Fraction of unpaired ssDNA bases as a function of force at different NaCl concentrations: 50 mM (green), 100 mM (blue), 250 mM (purple), 500 mM (orange), 1000 mM (yellow). Dashed lines are fits to Eq. 59. (**b**) Fraction of unpaired ssDNA bases as a function to Eq. 59. (**c**) Fraction of the unpaired bases at zero force $\phi(0)$ at different NaCl (black squares) and $MgCl_2$ (red diamonds) concentrations. Magnesium concentrations are multiplied by a factor 100. Lines are fits to the power law dependence, $\phi(0) \propto (c - c^*)^{-\gamma}$.

length; the so-called compensation effect between different structural motifs that unfold/fold independently due to thermal forces; and the low cooperativity of the unfolding/folding transition expected for such motifs.

In order to have a quantitative measure of the fraction of ssDNA that forms secondary structure, we used the WLC model to estimate the effective contour length L_c^{Eff} as a function of force. For each FEC we extracted the effective contour length by fixing the value of the persistence length to that obtained by fitting the elastic response in the range 10-40 pN. If the WLC model were to describe well the whole range of measured forces, the effective contour length would be equal to the one measured in the force range 10-40 pN (L_c^{10-40}) and should not depend on force even at low forces. The fraction of unpaired ssDNA bases is then simply given by the ratio $\phi(f) = \frac{L_c^{Eff}(f)}{L_c^{10-40}}$ (Figure 77**a, b**).

Below 50 mM NaCl, L_c^{Eff} increases at forces lower than 5 pN and the fraction $\phi(f)$ becomes greater than 1 (data not shown). This odd result indicates the inadequacy of the WLC model in this regime and is related to excluded volume effects that become prominent at low salt concentration. Above 100 mM NaCl, and in all explored concentrations of divalent salt, L_c^{Eff} decreases with decreasing force. We at-

Table 5: Parameters obtained from a fit to Eq. 59 of the curves presented in Figure 77**a,b**. For the NaCl case a fixed amount of 9 mM of salt concentration from the buffer must be added at each condition.

[NaCl] (mM)	f^* (pN)	δ (pN)	$\phi(0)$
100	0.440	2.090	0.448
250	3.6	1.75	0.313
500	4.850	1.637	0.049
1000	5.322	1.520	0.029
$[MgCl_2]$ (mM)	f^* (pN)	δ (pN)	$\phi(0)$
0.5	0.625	1.958	0.421
1	2.270	2.070	0.250
2	3.310	1.580	0.110
4	4.899	1.610	0.046
10	5.681	1.430	0.018

tribute this feature to the formation of secondary structure. At low forces (< 5 pN), the higher the salt concentration the shorter the effective contour length, meaning that at high monovalent salt concentration (or in presence of divalent salt) the fraction of paired bases increases. At the lowest measured forces (1-2 pN), the fraction of unpaired bases can reach up to 10-20% at the highest salt concentrations (namely 1 M of NaCl and 10 mM of $MgCl_2$).

In order to extract the fraction of unpaired bases at zero force, $\phi(0)$, we fit data to the following phenomenological formula,

$$\phi(f) = \frac{1}{1 + \exp\left[-\left(\frac{f - f^*}{\delta}\right)\right]} \tag{59}$$

where f^* is the critical force at which half of the nucleotides are paired while δ is a measure of the width of the force region in which ssDNA starts to form secondary structure. An analogous exponential dependence of the fraction of secondary structures has been suggested on the basis of mesoscopic models that take into account the formation of multiple and degenerate structures [310]. Fits to the experimental data using Eq. 59 are shown in Figure 77, and summarized in Table 5.

For both monovalent and divalent ions, f^* increases whereas the fraction of unpaired bases at zero force, $\phi(0)$, decreases with salt concentration, as expected. Figure 77**c** shows $\phi(0)$ as a function of salt concentration. At both 1 M of NaCl and 10 mM of $MgCl_2$ $\phi(0)$ has almost reached the minimal value of 0, meaning that at zero force and high salt almost all bases contribute to secondary structure formation. The dependence of $\phi(0)$ with salt concentration can be described by a power law $\propto (c - c^*)^{-\gamma}$ where c^* stands for a critical concentration and $\gamma > 0$ is an exponent. From the fit (Figure 77**c**) one gets $\gamma \approx 1$ for

both monovalent and divalent salt, but with a different value of c^* (65 mM and 0 mM for NaCl and $MgCl_2$). At high c we find $f^* > \delta$ meaning that we can approximate $\phi(0)$ in Eq. 59 by $\exp(-f^*/\delta)$ showing that f^* increases logarithmically with c, as indicated in the results of Table 5.

Both the elasticity measurements and the secondary structure formation in ssDNA shown in these sections rely on a trade-off between filament flexibility and electrostatics. Electrostatic repulsion between phosphate groups along the backbone stiffens the filament. Upon increasing salt concentration, the backbone self-repulsion is screened and hence its persistence length decreases. Moreover, the higher the salt, the larger the screening of the backbone, facilitating base pair formation. Overall, these effects contribute to the formation of non-native secondary structure. As a further step, it would be very interesting to apply this technique to study the FEC in presence other kind of cations both monovalent and divalent, such as potassium and calcium, or even multivalent cations such as hexamine cobalt compounds and polyamines. Finally, a deeper understanding of the kinetics and thermodynamics of non-specific secondary structure formation could be obtained by applying the present methodology to investigate ssDNA molecules of varying length and basepair composition.

CHAPTER B

MICROFLUIDIC SET-UP FOR DNA-LIGAND EXPERIMENTS IN THE MINITWEEZERS

This thesis focuses on the use of optical tweezers to investigate binding reactions between ligands (e.g. peptides, proteins, organic dyes) and nucleic acids. In many of such experiments one wants to tether a DNA/RNA molecule in ligand-free buffer, characterize the tethered molecule, and then flow the ligand into the microfluidic chamber. For instance, without this setup it would not be possible to perform the DNA-KF condensation experiments shown in section 14.1, as DNA molecules pre-incubated with the peptide form a strong condensate that cannot be tethered with the optical tweezers. Consequently, to investigate such binding reactions it is essential to exchange the buffer in the experimental set-up in a fast, but yet controlled way (i.e. without breaking the tethered molecule).

B.1 MICROFLUIDIC CHAMBER ASSEMBLY

The fluidics chambers used for this thesis (Figure 78) are assembled using two coverslips spaced with two nescofilm layers[1] that have been previously cut with a laser engraver with the appropriate channels design (Figure 78). An entry and exit hole for each channel are also drilled in one of the coverslips in order to flow in the buffer and beads. A glass micropipette and two dispenser tubes that connect the lateral channels to the central channel are embedded between the nescofilm layers, and the full assembly is heated for 5-10 min at 120°C to melt

1 Nescofilm (Sealing film). Karlan Research Products Corporation.

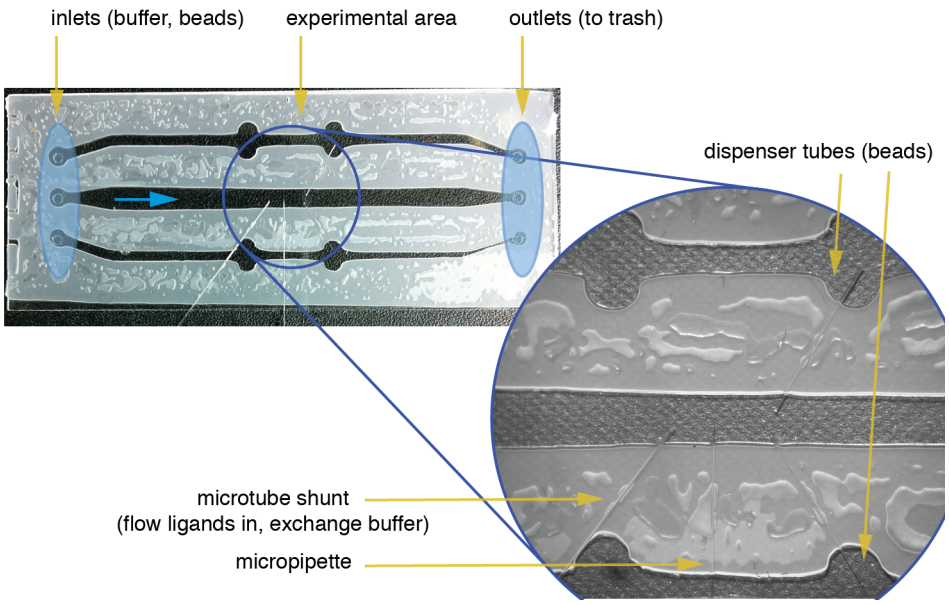

Figure 78: Microfluidic chamber. Picture of an assembled microfluidic chamber used for ligand-binding experiments. Beads are differentially flowed through lateral channels, and experiments are performed by subjecting one bead on an embedded micropipette and the other bead on the optic trap. Ligands can then be flowed in through a microtube shunt. The use of this delivery tubing directly placed on the experimental area reduces surface adsorption and time lag due to the chamber dead volume.

the nescofilm and stick everything together. Micropipettes are made using a self-made pipette-puller[2], and for most experiments a tip diameter of ~1 µm is used. The central channel has a volume of ~30 µl, and experiments are performed in it, near the tip of the micropipette. The two lateral channels are used to flow into the central channel beads or cells so they can be trapped with the lasers and experiments performed. Finally, for experiments in which ligands must be flowed into the chamber, a microtube shunt is also embedded[3] (Figure 78).

B.2 DELIVERY WITH MICROTUBE SHUNT OR TUBE JUNCTION

For most experiments shown in this thesis, a microtube shunt has been used to flow the ligands into the experimental area. The microtube shunt is connected through a capillary polyethylene tubing[4] to

2 A vertically subjected capillary glass tube (ID: 0.040(6) mm OD: 0.080(6) mm, King Precission Glass, Inc) is locally melted using a platinum wire with a ramping current protocol. A weight subjected to the end of the capillary tube pulls it creating the tip on the melting region. The size of the tip can be changed by adjusting the ramping speed.

3 Microtube shunt: ID: 0.080(6) mm OD: 0.13(1) mm. Bead dispenser tube: ID: 0.025(6) mm OD: 0.10(6) mm. King Precission Glass, Inc.

4 Polyethylene Tubing PE-10 (ID=0.28 mm, OD=0.61 mm). Intramedic™, Clay Adams™Brand.

the syringe that contains the ligand solution. The use of a microtube shunt to flow ligands in, is particularly useful as:

- Reduces the time lag required to fully equilibrate the chamber with the new buffer as the dead volume of the chamber and connecting tubing is greatly reduced (<20 µl for usual conditions[5]).
- Reduces surface adsorption of peptides, proteins and organic compounds, increasing the reproducibility of experiments at low concentrations.
- Unbinding kinetics experiments of highly hydrophobic ligands can be performed. This configuration allows to differentially flow ligands and ligand-free buffer through the microtube shunt or the central channel. This is specially useful for molecules with high adsorption to plastic and glass surfaces (e.g. Thiocoraline), as the washing out step can be performed through a clean non-saturated tubing.

A syringe pump (PicoPump, KDScientific) is generally used to flow ligand solutions[6]. For most experiments an air bubble is left at the end of the syringe in order to create a pressure reservoir that smoothens the flow, and flow-rates between 1-15 µl/min are used (causing a drag force of ~2-30 pN on a bead of 3 µm in diameter). To restrict and stop the buffer flow a pinch clamp is used to compress the PE-10 tubing[7].

Alternatively to the microtube shunt, buffers can also be flowed into the microfluidics chamber through the inlet of the central channel (Figure 78, blue). To do so, a PE-10 tubing and a PE-50 tubing are placed together into the silicone connector screw to create a double inlet. The junction is tightened together with parafilm or epoxy-glue, and in this way, buffer solutions can also be flowed into the chamber at a low flow-rate using the syringe pump. This method is less accurate and has a higher dead volume than the microtube shunt, but is especially useful to flow a "ligand-free" or scavenger buffer solution in unbinding kinetics experiments. In this way, the washing out step is performed without the potential interference of ligand desorption from the surfaces (relevant for highly hydrophobic compounds such as Thiocoraline, YOYO-1 or Kahalalide F).

B.3 MICROFLUIDIC CHAMBER COATING WITH mPEG

For experiments performed at subnanomolar ligand concentration or with proteins known to strongly bind to surfaces, it is useful to pas-

5 Typically, ~ 30 cm of PE-10 tubing are used to connect the microtube shunt to the syringe.

6 Although gravity force can also be used by uncapping the syringe and controlling its height.

7 The pinch clamp is assembled with a metallic support and an ultra-fine Hex adjuster screw (Thorlabs).

sivate the microfluidic chamber by coating it with mPEG (methoxy-polyethylene glycol). This creates an homogenous hydrophobic layer on the coverslip that strongly reduces nonspecific adsorption of proteins onto the glass surface[8]. The protocol used to coat the coverslip slides follows[9].

- Set coverslips on a glass staining jar for slides or in a coverslip Mini-Rack (Molecular Probes). Sonicate in acetone for 30 min. *On the meanwhile sonicate reaction containers for 20 min (two glass staining jars and a 250 ml erlenmeyer) with 1 M KOH.*
- Pour out acetone; sonicate coverslips in ethanol for 15 min. *Pour out KOH from reaction containers, rinse with MeOH, fill with MeOH and sonicate for 20 min.*
- Dispose ethanol, rinse coverslips with MilliQ water, and fill with 1 M KOH. Sonicate for <30 min[10]. *Rinse reaction containers with MeOH and let dry.*
- Dispose KOH, rinse coverslips with MilliQ water[11], and fill with Milli Q water. Take out coverslips, dry with N_2 and place in empty containers.
- Burn coverslips for each side with propane for ~ 1 sec to remove organic residues, and place in reaction containers.
- Prepare reaction solution in clean erlenmeyer (scale as needed): Pour 100 ml MeOH, add 5 ml glacial acetic acid and 1-1.5 ml APTES[12]. Mix well and pour solution in reaction containers.
- Incubate reaction for 10 min. Sonicate for 1 min and incubate again for 10 min.
- Rinse coverslips with MeOH and MilliQ water. Blow dry with N_2 and put in clean boxes. *For next use, rinse reaction flasks with MeOH, incubate for 20 min with 1M KOH and rinse with MilliQ water.*
- Make fresh buffer for pegylation: 84 mg sodium bicarbonate + 10 ml MilliQ water. Filter with 0.22 μm filter.

8 Specially if combined with other passivation/solubilization strategies (e.g. BSA, surfactants).

9 The glass capillary tubes used for the microtube shunt are also coated by flowing the solutions through a PE-10 tubing. The initial wash steps with acetone and KOH are omitted to avoid plastic tube melting.

10 Longer times might cause glass etching.

11 At this point a 45 min extra wash step with piranha solution (H_2O_2:H_2SO_4, 4:9) might be performed (dangerous, highly corrosive!). Equivalent results can be achieved using a plasma cleaner if available.

12 3-aminopropyltriethoxysilane, Sigma Aldrich. Store in vacuum/inert atmosphere.

- Disolve 80 mg mPEG aliquot[13] with 320 µl pegylation buffer. Mix gently by pippete or rack.
- Place half of the coverslips on top of pipette tip boxes filled with water, and drop ~70 µl pegylation solution on each slide. Place another coverslip gently on top of the slide sandwiching the solution (no bubbles). Incubate at dark for 4h to overnight.
- Disassemble sandwiched slides, rinse with MilliQ water, and store at dark (soaked in MilliQ water in containers sealed with parafilm; or in vacuum after blow drying with N_2). Use within 2 months.

Microfluidic chambers can then be assembled using the mPEG-coated coverslips and microtube shunts similarly to non-pegylated chambers. It is advisable however, to partly remove the mPEG coating from the sides of the coverslip to facilitate the assembly of the chamber, as the mPEG coating strongly reduces the stickiness of the glass surface. This is done by partially soaking the coated coverslips in 1 M KOH for <30 min. To do so a crystallizer beaker is filled with 1 M KOH up to a height of 0.5-1 cm. Coverslip slides are then left partially soaked into the beaker through its longest edge in order to remove the mPEG coating from the region that conforms the lateral channel and the nescofilm sticking area[14] (Figure 78). The coverslips are then soaked in MilliQ water and blow dried with N_2, and the same procedure is followed with the opposite side of the slide. In this way, an hydrophobic strap in the central region of the slide (along the central channel) can be created. The hydrophobic strap should be easily identifiable then by soaking the coverslips in MilliQ water.

13 mPEG-succinimidyl carboxymethyl MW: 5000 (ref: m-PEG-SCM-5000, Laysan Bio Inc.). Aliquots are stored at -20°C sealed with dissecant pellets. Take out aliquot from freezer and keep a few hours at dark and room temperature before use.

14 An old used chamber can be used as a reference.

CHAPTER C

METHODS OF COMPLEMENTARY TECHNIQUES

C.1 KAHALALIDE F AND THIOCORALINE SAMPLE PREPARATION

Stock solutions containing 2 mM KF were prepared by dissolving a ~1-2 mg sample of pure lyophilized KF (gently provided by Pharmamar) in 100% DMSO (Sigma, Molecular Biology Grade). Stock solutions were then vortexed at low speed for 20 min and filtered with a 0.2 µm pore PTFE membrane filter (Millipore) that was pre-rinsed with 300 µl DMSO to reduce filter extractables. Sample concentration after stock filtration was verified by means of HPLC: An aliquot before and after filtering was collected and stored in ACN:H_20 (50:50). These samples were sequentially processed in an HPLC using a 5%-100% acetonitrile gradient. The elution peak was monitored at 220 nm, and the integrated area was compared between both samples. No sample loss happened during filtration. Ready-to-use Kahalalide F stock solutions were stored at -20°C as 20-60 µl aliquots. To prepare diluted working solutions, a stock solution aliquot was thawed and sequentially diluted in TE 100 mM NaCl (*e.g.* 40 µM). The final DMSO concentration was always adjusted to 2%. Once thawed, stock solutions were not stored again to avoid the problems related to freeze-thaw cycles [311, 312]. Samples were always prepared in glass vials, and a glass syringe used for sample filtering. Use of glass material was preferred to avoid material leaching from disposable plasticware [313], and the good performance of glass infusion devices for low concentration samples of KF in clinical practice [314].

A similar approach was used to prepare aliquoted stock solutions of Thiocoraline.

C.2 MASS SPECTROMETRY METHODS

Sample are introduced using an Automated Nanoelectrospray. Triversa NanoMate (Advion BioSciences, Ithaca, NY, USA) sequentially aspirated the sample from a 384-well plate with disposable, conductive pipette tips, and infused it through the nanoESI Chip, which consists of 400 nozzles in a 20x20 array. Spray voltage was 1.75 kV and delivery pressure was 0.5 psi. Mass Spectrometer: Synapt HDMS (Waters, Manchester, UK). Samples were acquired with Masslynx software v.4 SCN 639 (Waters). **MS Conditions for TOF results:** NanoESI. Positive mode TOF. V mode. Sampling cone: 20 V. Source temperature: 20°C. Trap Collision Energy: 10. Transfer Collision Energy: 10. Trap Gas Flow: 8 ml/min. Vacuum Backing pressure: 5.89 mbar. m/z range: 300 to 5000, 500-15000. Instrument calibrated with CsI (external calibration). **MS Conditions for Ion Mobility results:** NanoESI. Positive mode Ion Mobility mode. V mode. Sampling cone: 20 V. Source temperature: 20°C. Trap Collision Energy: 10. Transfer Collision Energy: 10. Trap Gas Flow: 8 ml/min. IMS Wave Velocity: 300 m/s. IMS Wave Height: 9.5 V. Transfer Wave Velocity: 200 m/s. Transfer Wave Height: 8 V. Vacuum Backing pressure: 5.89 mbar. m/z range: 300 to 5000, 500-15000. Instrument calibrated with CsI (external calibration).

C.3 DYNAMIC LIGHT SCATTERING

C.3.1 *Principle of Dynamic Light Scattering*

Dynamic Light Scattering (also referred to as Photon Correlation Spectroscopy) is a technique that allows to determine the size of submicrometrical particles in a solution by measuring their Brownian motion (typical range of sizes: 1 nm-5 μm). Briefly, a laser light is passed through the solution containing the particles, and the intensity of the scattered light is recorded with a detector placed at 90°. When the laser beam travels through the solution, particles smaller than the laser wavelength (633 nm) cause Rayleigh scattering in all directions. The signal reaching the detector is the phase addition of light scattered by several particles, and therefore bright and dark spots will be seen across the detector due to constructive and destructive interference respectively. As the particles are submitted to thermal fluctuations, the intensity recorded in the detector will also fluctuate with time (as the relative distance between emitting particles changes), causing a speckle pattern. The time-scale relevant for such fluctuations is that in which particles move by one wavelength of light. From the Stokes-Einstein relation we know that for a spherical parti-

cle $D = \mu k_B T = \frac{k_B T}{6\pi\eta r}$ (where D is the diffusion coefficient, μ the mobility, η the medium viscosity and r the particle radius). For instance, for particles of 100 nm one would expect fluctuations on the intensity profile to occur on the order of $t \sim 50$ ms ($t \sim \langle \delta x^2 \rangle / D \sim \lambda^2 / D$).

From the previous explanation follows that electric field fluctuations in the detector are directly related to thermal fluctuations of the particles in the solution. Hence, the electric field autocorrelation function should be directly related to the particles Brownian motion:

$$g_1(\tau) = \frac{\langle E(t) \cdot E(t+\tau) \rangle}{\langle E(t) \rangle^2} \tag{60}$$

Which for a system undergoing Brownian motion corresponds to an exponential decay:

$$g_1(\tau) = e^{-\Gamma\tau} \tag{61}$$

However, experimentally, we have access to the fluctuations of the intensity on the detector (not the electric field). We can relate the intensity correlation function to the particle movement by taking into account that $I(t) = \|E(t)\|^2 = \sum_{m,n} E_m(t) \cdot E_n^*(t)$ (m, n being particle indexes), which is know as the Seigert relationship:

$$g_2(\tau) \equiv \frac{\langle I(t) \cdot I(t+\tau) \rangle}{\langle I(t) \rangle^2} = 1 + \|g_1(\tau)\|^2 \tag{62}$$

This expression being more usually found as:

$$g_2(\tau) = B \cdot [1 + \beta \|g_1(\tau)\|^2] \tag{63}$$

where B corresponds to a baseline and β to the instrumental response.

Consequently, by measuring the autocorrelation function of the light scattered by a solution is it possible to determine the diffusion coefficient of the particles in the sample, and if they are homogenous in size, its hydrodynamic radius.

C.3.2 *Dynamic Light Scattering Methods*

A Photon Correlation Spectrometer (PCS) 3D from LS INSTRUMENTS was used for DLS measurements. The instrument is equipped with a He-Ne laser (632.8 nm). Measurements of at least 90 s were recorded at an angle of 90°. The hydrodynamic radius was calculated by an exponential fitting of the first cumulant parameter. Standard deviations were calculated from the second cumulant. The measurement temperature of 25°C was maintained by a decaline bath, which matches the refractive index of glass and does not therefore interfere with the measurement. The evolution of the hydrodynamic radius was observed during 1 hour. The time indicated corresponds to the minutes

past after preparation of the sample and the beginning of the measurement.

Zeta potential measurements were carried out at 25°C with a Malvern Instrument Zetasizer Nano Z by laser Doppler electrophoresis. Disposable polystyrene cells were used. Solutions of peptide, DNA and the mixture were measured at the same concentration as the light scattering measurements were performed. Commercially available λ-DNA (New England Biolabs) was used for the experiments.

CHAPTER D

SYNTHESIS OF DNA/RNA MOLECULAR CONSTRUCTS

This appendix describes the methods used to synthesize the DNA and RNA molecular constructs for single-molecule manipulation used in this thesis. This includes molecules for dsDNA stretching experiments (section D.1); templates for DNA unzipping experiments (section D.2 and D.3), short DNA hairpins (section D.4 and D.5), and RNA motifs (section D.6). As a general remark oligonucleotides used for the synthesis of the final molecular constructs are always purchased purified by gel electrophoresis to avoid truncated products or oligonucleotides with misincorporated bases. This is specially important for the DNA hairpin synthesis, were long oligonucleotides (>40 nucleotides) are usually used.

D.1 24-kb dsDNA MOLECULE FOR STRETCHING EXPERIMENTS

For the DNA stretching experiments we prepared a 24508-bp DNA molecule with biotin and digoxigenin tags at the 3'-ends of the molecule (Figure 79). The DNA template was prepared by cleaving N6-methyladenine free λ-DNA (New England Biolabs) with XbaI restriction enzyme, and purified using the Wizard DNA clean-up system kit[1] (Promega).

1 This kit is well suited to purify DNA fragments >20 kb when using water or buffer pre-warmed at 80°C for elution. For the described synthesis the DNA was eluted in 50 µl Tris-Cl 10 mM, recovering ~170 ng/µl.

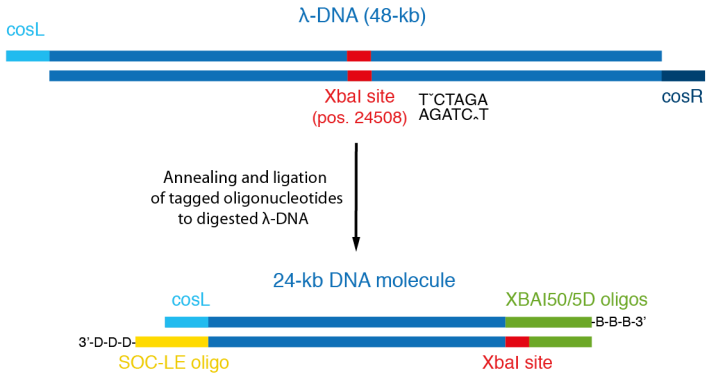

Figure 79: Scheme of the DNA stretching construct. Lambda-DNA (48-kb) is digested with XbaI creating two half λ-DNA molecules. The 24-kb fragment containing the cosL cohesive end (12-nt) is used to create the molecular construct for single-molecule manipulation. Oligonucleotides complementary to the cosL and XbaI cohesive ends are used to attach the digoxigenin and biotin tags respectively to each end of the molecule.

λ-DNA digestion:	
MilliQ water	62 µl
10X NEB reaction buffer 4	10 µl
100X BSA	1 µl
λ-DNA (0.5 µg/µl)	25 µl
XbaI enzyme (20 U/µl, NEB)	2 µl
Total volume:	100 µl

Incubate 3h at 37°C.

The digoxigenin tag was prepared by annealing to the cosL end of the molecule an oligonucleotide (SOC-LE) tailed at its 3'-end with digoxigenin-labeled dUTP's using terminal transferase (Roche). The biotin tag was prepared by annealing two complementary oligonucleotides designed to create an XbaI cohesive end. One of the oligonucleotides (XbaI50) was tailed at its 3'-end with multiple biotins with biotin-labeled dUTP's using terminal transferase (Roche). The oligonucleotides used for this synthesis are found in Table 6, and the protocol used for the tailing reactions is as follows:

Name	Oligonucleotide sequence
SOC-LE	5′-Pho-AGG TCG CCG CCC AAA AAA AAA AAA-3′
XbaI50	5′-Pho-CTA GAC CCG GGC TCG AGG ATC CCC-3′
XbaI5D	5′-GGG GAT CCT CGA GCC CGG GT-3′

Table 6: Oligonucleotides used for the synthesis of a DNA template for stretching experiments.

Oligonucleotide tailing reactions:	
MilliQ water	8 µl
5X $CoCl_2$ solution	4 µl
5X reaction buffer	4 µl
10 mM dATP	1 µl
1 mM DIG-dUTP or BIO-dUTP[2]	1 µl
100 µM oligonucleotide (SOC-LE or XbaI50)	1 µl
Terminal transferase (400U/µl)	1 µl
Total volume:	20 µl
Incubate 15′ at 37°C. Add 1 µl EDTA 0.5 M to quench reaction.	

Oligonucleotides were purified after tailing steps using the QIAquick Nucleotide Removal Kit (Qiagen) and eluted in 50 µl Tris-Cl 10 mM. The annealing of the oligonucleotides to the half λ-DNA was then performed in two steps. First the two oligonucleotides that conform the right-arm (XbaI50-biotinylated and XbaI5D) were annealed using the following protocol:

Annealing of right-arm (biotin tag):	
oligo XabI50 biotin-tailed (~2 µM)	48 µl
oligo XbaI5D (100 µM)	1 µl
Annealing buffer	5 µl
Total volume:	54 µl
Annealing buffer:	
1 M Tris pH7.5	100 µl (100 mM)
1 M $MgCl_2$	10 µl (10 mM)
MilliQ water	890 µl
Total volume:	1000 µl
Protocol: 95°C for 1′; cool down from 80°C to 10°C in steps of 0.5°C every 10″	

To anneal the final molecular construct, first the digoxigenin-tailed oligo (SOC-LE) was annealed to the half λ-DNA molecule by incu-

2 Digoxigenin-11-dUTP 1 mM, alkali-stable; and Biotin-16-dUTP, tetralithium salt, 1 mM (Roche).

bation for 10 min at 68°C in a 10-fold excess oligonucleotide. Then, the biotin tag (right-arm) was annealed by incubation for 1h at 42°C (using 20-fold excess oligonucleotides) followed by cooling down to room temperature.

Annealing of molecular construct:	
digested λ-DNA (170 ng/µl)	40 µl
oligo SOC-LE DIG-tailed (~2 µM)	5 µl
Annealing buffer	5 µl
Total volume:	50 µl

Protocol: Incubate 10′ at 72°C. Place in thermal bath at 42°C, and add 10 µl of right-arm construct. Incubate for 1h at 42°C and let cool down to room temperature.

Finally the annealed construct is ligated using T4 DNA ligase in an overnight reaction:

Ligation reaction:	
MilliQ water	7 µl
Annealing product	50 µl
10X T4 DNA ligase buffer	7 µl
10 mM ATP	3 µl
T4 DNA ligase (400 U/µl)	3 µl
Total volume:	70 µl

Protocol: Overnight reaction at 16°C.

The ligated construct is then thermally inactivated by incubation for 10′ at 65°C and aliquoted for storage at -20°C. 1-2 µl EDTA 0.5 M can also be added for storage.

D.2 6.7-KB DNA HAIRPIN WITH SHORT HANDLES (29 BP)

For the DNA unzipping experiments a 6838-bp DNA hairpin with a tetraloop at one end and two 29-bp dsDNA handles at the other end was prepared. The synthesis is based on a previously described method [62], but was modified to introduce double digoxigenin and biotin tags at each handle to enhance tether lifetimes (Figure 80). The sequence of the oligonucleotides used for the hairpin synthesis, together with the 30-base oligonucleotide used to generate the ssDNA template are specified in Table 7. To create a 6.8 kb DNA hairpin, N6-methyladenine free λ-DNA (New England Biolabs) is digested with the restriction enzyme BamHI, and the fragment contained between positions 41733 and 48502 (cosR end) used as the stem of the hairpin.

λ-*DNA digestion:*	
MilliQ water	123 µl
10X NEB reaction buffer 3	20 µl
100X BSA	2 µl
λ-DNA (0.5 µg/µl)	50 µl
BamHI enzyme (20 U/µl, NEB)	5 µl
Total volume:	200 µl

Incubate 4h at 37°C.

The digested λ-DNA is then purified using the Wizard SV Gel and PCR clean-up system (Promega) or equivalent (2 columns, elution in 50 µl Tris-Cl 2 mM), and phosphorylated[3] with T4 polynucleotide kinase (New England Biolabs).

λ-*DNA phosphorilation:*	
MilliQ water	4 µl
10X PNK buffer	10 µl
10 mM ATP	10 µl
digested λ-DNA (~180 ng/µ;)	75 µl
Total volume:	100 µl

Incubate 30′ at 37°C.

After phosphorylation, the digested λ-DNA is purified with the Wizard SV Gel and PCR clean-up system (1 column, elution in 50 µl Tris-Cl 2 mM), and is ready for annealing with the oligonucleotides that create the loop and handles structure. To create the end-loop of the molecule, the stem is annealed to an oligonucleotide (BamHI-loop2) that self-assembles into a hairpin structure with a tetraloop at one end, and an overhang complementary to the BamHI restriction site on the other end. To create the dsDNA handles, we used two partially complementary oligonucleotides (cosRlong and Bio-cosRshort3) that hybridize forming a protruding end complementary to the cosR end. The Bio-cosRshort3 oligonucleotide was purchased 5′-biotinylated, and the cosRlong oligonucleotide was tailed with multiple digoxigenins at its 3′-end with digoxigenin-labeled dUTP's using terminal transferase (Roche). To create the doubly biotynilated dsDNA handle, a third oligonucleotide (splint3) complementary to the unpaired regions of BIO-cosRshort3 was also tailed at its 3′ end with multiple biotins with biotin-labeled dUTP's. To create the digoxigenin dsDNA handle we used a modified oligonucleotide (inverted-splint, Thermo Scientific) complementary to the unpaired region of cosRlong. This

3 Although the end digested by BamHI is fully phosphorylated, the cosR end of λ-DNA is not.

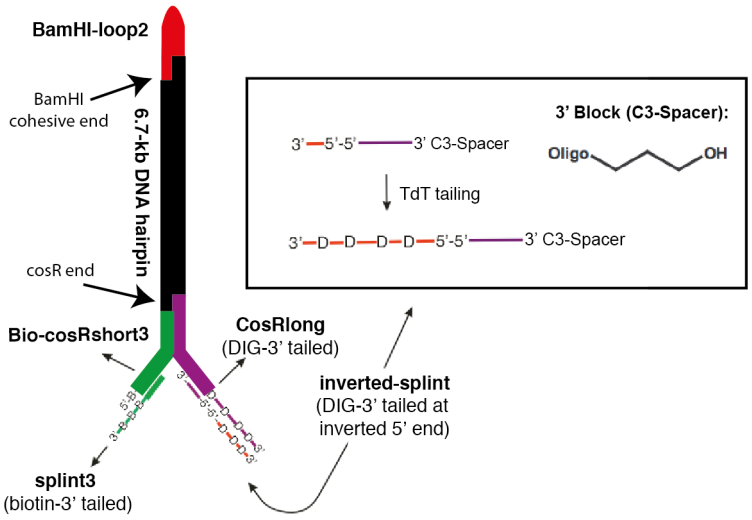

Figure 80: Scheme of the 6.7-kb hairpin synthesis. The DNA hairpin is created by ligating a set of oligonucleotides (red, green, purple) to a 6.7-kb restriction fragment of λ-DNA (black). The cosR end, and an BamHI cohesive end were respectively used to anneal the dsDNA handles and the end-loop to the λ-DNA fragment. To create the biotinylated handle (green) one oligonucleotide was purchased 5′-biotinlylated (Bio-cosRshort3) and the other one was tailed with multiple biotins at its 3′ end (splint3). To create a dsDNA handle with multiple digoxigenins we used an oligonucleotide containing a 5′-5′ inversion and a blocked 3′ end (inverted-splint). In this way both oligonucleotides (cosRlong, inverted-splint) were tailed with multiple digoxigenins at the appropriate end.

oligonucleotide contains two modifications: a C3 spacer at its 3′ end to block this end in tailing reactions, and a polarity inversion at its 5′ end using a 5′-5′ linkage. This end could therefore be tailed with digoxigenin-labeled dUTP's using the same tailing protocol. In this way, both ends of the handle could be tailed with multiple digoxigenins (Figure 80).

Oligonucleotide tailing reactions:	
MilliQ water	8 µl
5X $CoCl_2$ solution	4 µl
5X reaction buffer	4 µl
10 mM dATP	1 µl
1 mM DIG-dUTP or BIO-dUTP[4]	1 µl
100 µM oligonucleotide	1 µl
Terminal transferase (400U/µl)	1 µl
Total volume:	20 µl

Incubate 15′ at 37°C. Add 1 µl EDTA 0.5 M to quench reaction.

4 Digoxigenin-11-dUTP 1 mM, alkali-stable; and Biotin-16-dUTP, tetralithium salt, 1 mM (Roche).

Name	Oligonucleotide sequence
BamHI-loop II	5′-Pho-GAT CGC CAG TTC GCG TTC GCC AGC ATC CGA CTA CGG ATG CTG GCG AAC GCG AAC TGG C-3′
cosRlong	5′-Pho-GGG CGG CGA CCT AAG ATC TAT TAT ATA TGT GTC TCT ATT AGT TAG TGG TGG AAA CAC AGT GCC AGC GC-3′
bio-cosRshort3	5′-Bio-GAC TTC ACT AAT ACG ACT CAC TAT AGG GAA ATA GAG ACA CAT ATA TAA TAG ATC TT-3′
splint3	5′-TCC CTA TAG TGA GTC GTA TTA GTG AAG TC-3′
inverted-splint	3′-AAA AA-5′-5′-GCG CTG GCA CTG TGT TTC CAC CAC TAA C(SpC3)-3′
Blockloop30	5′-TAG TCG GAT GCT GGC GAA CGC GAA CTG GCG-3′

Table 7: Oligonucleotides used for the DNA hairpin synthesis and to generate the ssDNA template.

Oligonucleotides were purified after tailing steps using the QIAquick Nucleotide Removal Kit (Qiagen) and eluted in 50 µl Tris-Cl 10 mM. The final molecular construct is then assembled with a ~10-fold excess handles and 100-fold excess loop:

Annealing reaction:	
MilliQ water	7.3 µl
digested λ-DNA (330 ng/µl)	29 µl (0.3 pmol)
oligo cosRlong dig-tailed (2 µM)	5 µl (10 pmol)
oligo BamHI-loop II (5 µM)	5 µl (25 pmol)
oligo inverted-splint dig-tailed (2 µM)	5 µl (10 pmol)
oligo splint-3 biotin-tailed (2 µM)	5 µl (10 pmol)
oligo bio-cosRshort (5 µM)	1 µl (5 pmol)
1 M Tris pH7.5	1.5 µl (25 mM)
5 M NaCl	1.2 µl (100 mM)
Total volume:	60 µl
Protocol: 70°C for 10′; 55°C for 10°C and let cool down to room temperature in thermal bath with cooling on.	

Finally the annealed construct is ligated using T4 DNA ligase in an overnight reaction[5].

5 An optional microdialysis in a low salt buffer (Tris-Cl 10 mM, EDTA 1 mM, NaCl 10 mM, pH7.5) can be done if the final NaCl concentration in the annealed construct is >100 mM, and the sample can not be diluted before ligation.

Ligation reaction:	
MilliQ water	9 µl
Annealing product	55 µl
10X T4 DNA ligase buffer	8 µl
10 mM ATP	5 µl
T4 DNA ligase (400 U/µl)	3 µl
Total volume:	80 µl

Protocol: Overnight reaction at 16°C.

The ligated hairpin is then thermally inactivated by incubation for 10′ at 65°C and aliquoted for storage at -20°C[6].

D.3 SYNTHESIS OF LONG DNA HAIRPINS (>40 BP) OF ARBITRARY SEQUENCE

This method is used to synthesize the DNA hairpin for the single molecule footprinting experiments. However the protocol has been designed so it can be directly used to synthesize hairpins of arbitrary sequence and length as far as they do not contain the restriction sites of TspRI and Tsp45I in its sequence[7].

A DNA sequence of interest is designed by adding to each end of the desired sequence the restriction sites of TspRI and Tsp45I flanked by two primer binding sites optimized for PCR extension of the region (Figure 81**a**). The two restriction enzymes have the following degenerate restriction sites (5′-$_{\uparrow}$NNCASTGNN$^{\downarrow}$-3′ for *TspRI* and 5′-$^{\downarrow}$GTSAC$_{\uparrow}$-3′ for *Tsp45I*). The enzymes have been selected to create long cohesive ends that permit an efficient annealing and ligation of the DNA handles and end-loop after restriction (Figure 81**b**). The specific recognition sequences used for each enzyme (5′-CGCACTGAC-3′ 5′-GTCAC-3′) have also been chosen to be non-palindromic (to avoid self-complementarity) and to maximize duplex stability at the annealing region.

A standard vector containing the sequence of interest is then purchased from a provider of genomic services (e.g. Eurofins MWG Operon, Integrated DNA technologies)[8], and PCR amplified using the designed flanking primers (Table 8).

6 Instead of a thermal inactivation, a micro dialysis in low salt buffer (Tris-Cl 10 mM, EDTA 1 mM, NaCl 10 mM, pH7.5) can also be done.

7 The method can be extended to any sequence by replacing the selected enzymes by equivalent ones (> 4 nt overhang) not present in the template, or by modifying the sequence to remove the inner binding sites

8 The plasmid can also be synthesized as described in section D.6.1, albeit likely at a higher cost.

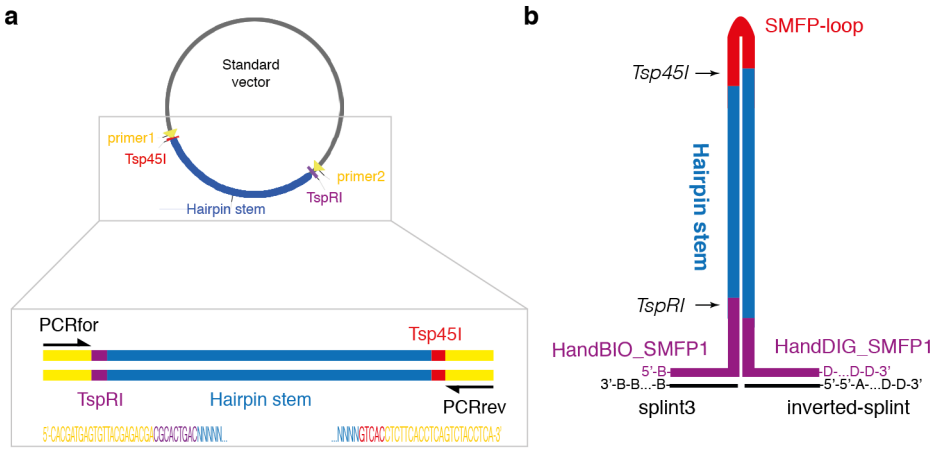

Figure 81: Scheme of the DNA hairpin synthesis. (**a**) The sequence that will become the stem of the hairpin (blue) is placed between the restriction sites of TspRI (purple) and Tsp45I (red) and two primer binding sites (yellow). This region of interest is purchased cloned into a standard commercial vector. (**b**) The DNA hairpin is created by ligating a set of oligonucleotides (red, purple, black) to the PCR-amplified and digested fragment from panel a. The oligonucleotides that conform the handles are differentially end-labelled with digoxigenins and biotins at both strands using the same approach than in section D.2.

PCR reaction conditions (1X):	
MilliQ water	31 µl
10X KOD buffer #1	5 µl
2 mM dNTPs	5 µl
25 mM $MgSO_4$	2 µl
10 µM primer SMfootpr-PCRfor	2.5 µl
10 µM primer SMfootpr-PCRrev	2.5 µl
1-10 ng/µl plasmid DNA	1 µl
KOD polymerase	1 µl
Total volume:	50 µl

PCR Procotol: 95°C for 30"; 25 x (95°C for 15", 58°C for 2", 72°C for 20")
(for a hairpin <1 kbp). Prepare reaction on ice and place as hot start (preheat at 95°C).

A total volume of 5X PCR reactions was performed (scale as needed), and the correct size of the product was checked in an analytic gel electrophoresis. The amplification product was then purified using the QIAquick PCR purification kit, and eluted in 50 µl Tris-Cl 10 mM, recovering ~425 ng/µl. The purified DNA was then digested with the two restriction enzymes:

Double digestion reaction:	
MilliQ water	41 µl
PCR fragment (~425 ng/µl)	40 µl
10X NEB buffer 4 (or CutSmart)	10 µl
100X BSA	1 µl
TspRI (10 U/µl)	3 µl
Tsp45I (5U/µl)	5 µl
Total volume:	100 µl

Incubate 4h at 65°C.

(Tsp45I is not well-suited for extended digestion, so further enzyme might be added after 2h.)

The digested DNA fragment (416 bp for the footprinting template) is then gel purified (2% agarose gel, 1X TBE) using Sybr Safe staining and blue illumination to minimize DNA damage and nicking. The DNA fragment is extracted from the gel using the QIAquick gel extraction kit (~80 ng/µl eluted in ~100 µl Tris-Cl 10 mM, two columns).

To assemble the final construct, an oligonucleotide that folds into a hairpin structure with a tetraloop at one end and a Tsp45I overhang at the other end is purchased (Figure 81 and Table 8). To create the DNA handles two oligonucleotides that are partially complementary and anneal creating a TspRI cohesive end are purchased (HandBIO_SMFP1 and HandDIG_SMFP1 in Table 8). To create the handles with multiple digoxigenin and biotin tags at both strands, we used the same approach and "splint" oligonucleotides described in section D.2. Briefly, to create the doubly biotyinylated dsDNA handle, a third oligonucleotide (splint3) complementary to the unpaired region of HandBIO_SMFP1 was also purchased. To create the digoxi-

Name	Oligonucleotide sequence
SMfootpr-PCRfor	5′-CAC GAT GAG TGT TAC GAG ACG A-3′
SMfootpr-PCRrev	5′-TGA GGT AGA CTG AGG TGA AGA G-3′
SMFP-loop	5′-Pho-GTC ACT TAG TAA CTA ACA TGA TAG TTA CTT TTG TAA CTA TCA TGT TAG TTA CTA A-3′
HandDIG_SMFP1	5′-Pho-AAG ATC TAT TAT ATA TGT GTC TCT ATT AGT TAG TGG TGG AAA CAC AGT GCC AGC GC-3′
HandBio_SMFP1	5′-Bio-GAC TTC ACT AAT ACG ACT CAC TAT AGG GAA ATA GAG ACA CAT ATA TAA TAG ATC TTC GCA CTG AC-3′
splint3	5′-TCC CTA TAG TGA GTC GTA TTA GTG AAG TC-3′
inverted-splint	3′-AAA AA-5′-5′-GCG CTG GCA CTG TGT TTC CAC CAC TAA C(SpC3)-3′

Table 8: Oligonucleotides used for the DNA hairpin synthesis.

genin dsDNA handle we used a modified oligonucleotide (inverted-splint, Thermo Scientific) complementary to the unpaired region of HandDIG_SMFP1. This oligonucleotide contains two modifications: a C3 spacer at its 3′ end to block this end in tailing reactions, and a polarity inversion at its 5′ end using a 5′-5′ linkage. In this way, both ends of the handle could be tailed with multiple digoxigenins. The oligonucleotides HandDIG_SMFP1 and inverted-splint are then tailed with digoxigenin-labeled dUTP's, and HandBIO_SMFP1 with biotin-labeled dUTP's using the following protocol:

Oligonucleotide tailing reactions:	
MilliQ water	8 µl
5X $CoCl_2$ solution	4 µl
5X reaction buffer	4 µl
10 mM dATP	1 µl
1 mM DIG-dUTP or BIO-dUTP[9]	1 µl
100 µM oligonucleotide	1 µl
Terminal transferase (400U/µl)	1 µl
Total volume:	20 µl

Incubate 15′ at 37°C. Add 1 µl EDTA 0.5 M to quench reaction.

Oligonucleotides were purified after tailing steps using the QIAquick Nucleotide Removal Kit (Qiagen) and eluted in 50 µl Tris-Cl 10 mM. The final molecular construct is then assembled with a ~10-fold excess handles and 100-fold excess loop:

Annealing reaction:	
416-bp digested PCR fragment (79.5 ng/µl)	4 µl (1.2 pmol)
oligo splint-3 biotinylated (2 µM)	6 µl (12 pmol)
oligo inverted-splint dig-tailed (2 µM)	6 µl (12 pmol)
oligo HandDIG_SMFP1 dig-tailed (2 µM)	6 µl (12 pmol)
oligo HandBIO_SMFP1 (10 µM)	1 µl (10 pmol)
oligo SMFP_loop (100 µM)	1 µl (100 pmol)
1 M Tris pH7.5	0.5 µl (20 mM)
5 M NaCl	0.5 µl (100 mM)
Total volume:	25 µl

Protocol: 70°C for 5′; cool down to 10°C in steps of 0.5°C every 20″

The assembly is then ligated in an overnight reaction using T4 DNA ligase (NEB) at 16°C.

9 Digoxigenin-11-dUTP 1 mM, alkali-stable; and Biotin-16-dUTP, tetralithium salt, 1 mM (Roche).

Ligation reaction:	
Annealing product[10]	25 µl
MilliQ water	5 µl
10X T4 DNA ligase buffer	4 µl
10 mM ATP	1.5 µl
T4 DNA ligase (400 U/µl)	2.5 µl
Total volume:	40 µl

Protocol: Overnight reaction at 16°C.

Finally the ligated hairpin is gel purified to remove the excess oligonucleotides (1.8%-2% agarose gel in 1X TBE) using Sybr Safe staining and blue led illumination. The hairpin is recovererd using the QIAquick Gel extraction kit and eluted in 50 µl Tris-Cl 10 mM. Finally, 0.5 µl 1M NaCl and 0.5 µl 100 mM EDTA are added to the sample for long term storage of the hairpin (EDTA 1 mM, NaCl 10 mM). The sample is usually aliquoted and stored at -20°C[11].

D.4 SHORT DNA HAIRPINS ($\leqslant$ 40 BP) WITH SHORT HANDLES (29 BP)

To synthesize short DNA hairpins with 29-bp dsDNA handles, the desired DNA sequence is purchased in a series of oligonucleotides that self-assemble onto the hairpin+handles structure (Figure 82). Briefly, a standard sequence is used to create the handles, and the hairpin sequence is embedded between them (Table 9). For very short hairpins (<13 bp) the structure can be ordered as a single oligonucleotide; however, for longer hairpins it is useful to split the molecular construct into two oligonucleotides that are annealed and ligated together. Finally, the dsDNA handles are created by annealing a third oligonucleotide that is complementary to the handle region ("splint"). The sequence of a series of oligonucleotides used to create some of the DNA hairpins used in this thesis are shown in Table 9. A guideline on the criteria used to split the molecular construct between two oligonucleotides at an appropriate position of the hairpin is found in section D.5.2.

For short hairpins (single oligo construct), the oligo is purchased 5'-biotinylated and tailed at its 3'-end with multiple digoxigenins using the DIG oligonucleotide tailing kit (Roche). For longer hair-

11 The gel and column purification of the construct might dissociate the 29-bp dsDNA handles. For the purpose of our experiments we did not observe a decreased stability of the hairpin tethers. However, one could ensure the presence of the "splints" in the final construct by performing a final annealing step with additional splint oligonucleotides (e.g. equimolar conditions). This should not severely increase the stickiness of the beads, as the splints can not self-assemble and create short tethers between beads.

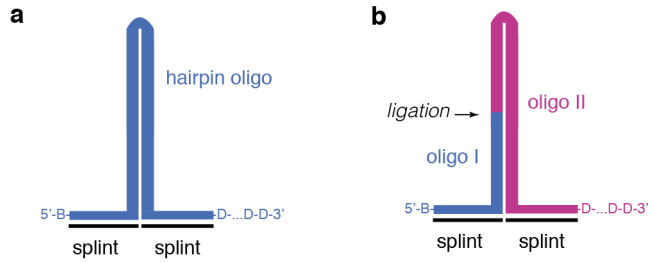

Figure 82: Scheme of short hairpin synthesis. Short hairpins are performed by annealing a set of complementary oligonucleotides. (**a**) For short sequences (stem < 13 bp), a single oligonucleotide is annealed to the splints that conform the dsDNA handles. (**b**) For longer stems, the sequence is split at an appropriate position between two oligonucleotides.

Name	Oligonucleotide sequence
EchiXS	5′-Bio-*AGT TAG TGG TGG AAA CAC AGT GCC AGC GC***C CCT ACG TAT ATG AAA ATA TAC GTA GGG** *AGT TAG TGG TGG AAA CAC AGT GCC AGC GC*-3′
EchiS-I	5′-Bio-*AGT TAG TGG TGG AAA CAC AGT GCC AGC GC***C CCT ACG**-3′
EchiS-II	5′-Pho-**TAC GTA TAG AAA TAT ACG TAC GTA GGG***AGT TAG TGG TGG AAA CAC AGT GCC AGC GC*-3′
splint	5′-GCG CTG GCA CTG TGT TTC CAC CAC TAA CT-3′

Table 9: Oligonucleotides used for the synthesis of two typical DNA hairpins: one that can be purchased as a single oligonucleotide (EchiXS), and a longer one that must be split (EchiS). The hairpin sequence and the handle sequence are highlighted in bold and italics respectively. A standard sequence is used for the handles, and so a single "splint" oligonucleotide can be used for all constructs.

pins (two oligos construct), a similar approach is used with the two oligonucleotides that create the structure: one oligo is purchased 5'-biotinylated, whereas the oligo containing the opposite handle is purchased 5'-phosphorylated and then tailed at its 3'-end with multiple digoxigenins. In this way, the two oligos can be assembled and ligated into the full molecular construct. The tailing reaction is performed as in previous protocols:

Oligonucleotide tailing reactions:	
MilliQ water	8 µl
5X $CoCl_2$ solution	4 µl
5X reaction buffer	4 µl
10 mM dATP	1 µl
1 mM DIG-dUTP or BIO-dUTP	1 µl
100 µM oligonucleotide (e.g. EchiXS, EchiS-II)	1 µl
Terminal transferase (400U/µl)	1 µl
Total volume:	20 µl
Incubate 15' at 37°C. Add 1 µl EDTA 0.5 M to quench reaction.	

After tailing steps oligos are purified using the QIAquick Nucleotide Removal Kit (Qiagen) and eluted in 50 µl Tris-Cl 10 mM. The final construct is then assembled in an equimolar reaction. Annealing is performed in one step for single oligonucleotide hairpins (oligo+splint) and in two steps for hairpins split in two oligonucleotides (oligoI+splint, oligoII+splint). The annealing protocol for both types of hairpins follows:

Annealing reaction for single oligonucleotide constructs:	
oligo dig-tailed (e.g. EchiXS) (~ 2 µM)	5 µl
oligo splint (5 µM)	2 µl
1 M Tris	1 µl
5 M NaCl	1 µl
MilliQ water	21 µl
Total volume:	30 µl
Protocol: Incubate each reaction at 95°C for 1', 80°C for 1', decrease by 0.5°C every 10" down to 10°C.	

Annealing reaction for two oligonucleotide constructs		
oligo I (e.g. EchiS-I) (5 µM)	1 µl	-
oligo II dig-tailed (e.g. EchiS-II) (~2 µM)	-	5 µl
oligo splint (5 µM)	1 µ	1 µl
1 M Tris	0.5 µl	0.5 µl
5 M NaCl	0.5 µl	0.5 µl
MilliQ water	12 µl	8 µl
Total volume:	15 µl	15 µl

Protocol: Incubate each reaction at 95°C for 1', 80°C for 10", decrease by 0.5°C every 10" down to 40°C. Hold temperature, mix both reactions and decrease by 0.5°C every 20" down to 10°C.

The assembly is then ligated in an overnight reaction using T4 DNA ligase (NEB) at 16°C.

Ligation reaction:	
Annealing product[12]	30 µl
MilliQ water	52 µl
10X T4 DNA ligase buffer	10 µl
10 mM ATP	5 µl
T4 DNA ligase (400 U/µl)	3 µl
Total volume:	100 µl

Protocol: Overnight reaction at 16°C.

D.5 SHORT DNA HAIRPINS (⩽ 40 BP) WITH LONG HANDLES (500BP)

To synthesize short DNA hairpins with long handles the following strategy is used (Figure 83):

- Synthesize each DNA handle by PCR reaction of a ~700 bp region of a plasmid that contains the restriction site of a non-palindromic enzyme with long cohesive ends (>4 nucleotides). Use a different enzyme for each handle, and place the restriction site at an asymmetric position of the fragment.
- Tail each PCR product with the appropriate tethering agent (biotin, digoxigenin), digest with the restriction enzyme, and gel purify the band that corresponds to the handle.

12 The ratio final volume/annealing volume is relevant to maintain the salt concentration low. A high salt concentration decreases ligation efficiency.

Name	Oligo/primer sequence
HandleAIII-for	5′-GCT TCC TAA TGC AGG AGT CG-3′
HandleAIII-rev	5′-GCA GAT CCG GAA CAT AAT GG-3′
B-Bio	5′-Bio-TTCTTGAAGACGAAAGGGCCT-3′
B-pstI	5′-CCATTGCTGCAGGCATCGTG-3′
EcoRIlong-I	5′-Pho-**GTC AC**C CGC GTC TTC AAG AAT TC-3′
EcoRIlong-IIB	5′-Pho-TCA TGT TTG TCC GAA AGG ACA AAC ATG AGA ATT CTT GAA GAC GCG G**GT CAC TGG T**-3′

Table 10: Primers used for the amplification of the handles, and an example of two oligonucleotides used to create a DNA hairpin for EcoRI binding experiments. In bold the cohesive ends for annealing to the handles.

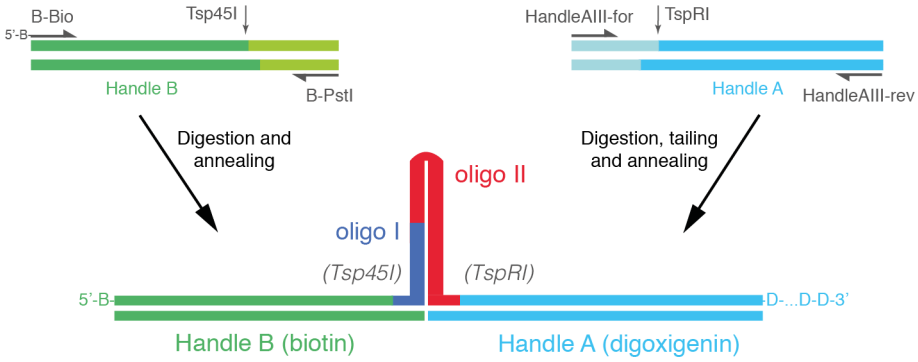

Figure 83: Scheme of hairpins with long handles. Long handles (~500 bp) are performed by PCR amplification of regions of plasmid pBR322 that contain a restriction site for TspRI/Tsp45I. Biotin and digoxigenin tags are introduced on the PCR primers or with a tailing reaction. After digestion, the handles are annealed and ligated to the oligonucleotides that fold into the hairpin structure.

- Purchase two oligonucleotides that self-assemble into the desired DNA hairpin and have cohesive ends complementary to the DNA handles.
- Anneal and ligate the oligonucleotides and handles.

D.5.1 *Synthesis of long handles*

Handles are performed by PCR amplification of plasmid pBR322. The PCR for Handle A is performed between positions 618 to 1674 (1057 bp).

PCR Handle A (1X):	
MilliQ water	45 µl
10X Biotaq buffer #1	10 µl
2 mM dNTPs	10 µl
10 µM primer HandleAIII-for	10 µl
10 µM primer HandleAIII-rev	10 µl
50 mM $MgCl_2$	4µl
1 ng/µl pBR322/SphI	10 µl
Biotaq polymerase	1 µl
Total volume:	100 µl

PCR Procotol: 95°C for 1'; 30 x (95°C for 45", 59°C for 1', 72°C for 1'30"); 72°C for 5'

The PCR-amplified fragment (8X reaction) is purified using the Wizard SV gel and PCR clean-up system and eluted in 50 µl Tris-Cl 2mM (~ 600 ng/µl). The purified PCR product is tailed using the 3'-5' exonuclease activity of T4 DNA polymerase:

Handle A tailing reaction:	
MilliQ water	39 µl
purified Handle A (~600 ng/µl)	35 µl
T4 DNA polymerase buffer (NEB buffer 2)	10 µl
BSA 100X	1 µl
1 mM digoxigenin-11-dUTP	10 µl
T4 DNA polymerase enzyme	5 µl
Total volume:	100 µl

Incubate at room temperature for 20 min.

The tailed DNA fragment is then purified using the Wizard SV gel and PCR clean-up system (elution in 50 µl Tris-Cl 2mM, ~ 270 ng/mul), and digested with TspRI enzyme. The fragment contains a single site for the enzyme at position 890 (5'-GTCACTGGT-3').

TspRI digestion reaction:	
MilliQ water	2.5 µl
tailed Handle A (~270 ng/µl)	40 µl
10X NEB buffer IV	5 µl
100X BSA	0.5 µl
TspRI (10 U/µl)	2 µl
Total volume:	50 µl

Incubate 4h at 65°C.

The digestion product is gel purified (2% agarose gel, 1X TBE) using Sybr Safe staining and blue illumination to minimize DNA damage and nicking. The longest fragment (784-bp) is excised and recovered using Qiaquick gel extraction kit (~115 ng/µl). This 784-bp fragment that contains a TspRI compatible end, will be used as handle A.

Handle B is performed by PCR amplification of the plasmid between positions 3601-4361 (761 bp), and using a 5'-biotinylated primer (B-Bio).

PCR Handle B (1X):	
MilliQ water	49 µl
10X Accuzyme buffer #1	10 µl
2 mM dNTPs	10 µl
10 µM primer B-Bio	10 µl
10 µM primer B-PstI	10 µl
1 ng/µl pBR322/SphI	10 µl
Accuzyme polymerase (or Taq)	1 µl
Total volume:	100 µl

PCR Procotol: 95°C for 1'; 30 x (95°C for 45", 56°C for 1', 72°C for 1'30"); 72°C for 5'.

The PCR-amplified fragment (9X reaction) is purified using the Wizard SV gel and PCR clean-up system, eluted in 50 µl Tris-Cl 2mM (~ 650 ng/µl), and digested with Tsp45I enzyme. The fragment contains a site for the enzyme at position 3834 (5'-GTCAC-3').

Tsp45I digestion reaction:	
MilliQ water	18 µl
Handle B (~650 ng/µl)	40 µl
10X CutSmart Buffer	7 µl
Tsp45I (5 U/µl)	5 µl
Total volume:	70 µl

Incubate 4h at 65°C.

The digestion product is gel purified (2% agarose gel, 1XTBE) using Sybr Safe staining and blue illumination to minimize DNA damage and nicking. The longest fragment (528-bp) is excised and recovered using Qiaquick gel extraction kit (~210 ng/µl in 50 µl 10 mM Tris-Cl). This 528-bp DNA fragment that contains a Tsp45I compatible end, will be used as handle B.

D.5.2 *Design of hairpin oligonucleotides*

To create the DNA hairpin, two oligonucleotides that self-assemble into the desired DNA hairpin and have overhangs complementary to Tsp45I and TspRI cohesive ends are designed and purchased (Table 10). The following guidelines are useful to split the hairpin into the two oligonucleotides at an appropriate position:

- Split the hairpin at at location at least 9-10 bp below the loop position in order to facilitate loop formation in the oligonucleotide that contains it. Check melting temperature of the self-assembling region.
- Split the hairpin at at location at least 6-7 bp above the stem start in order to facilitate annealing of the two oligonucleotides. Check melting temperature of the annealing region.
- Use M-fold to check that the oligonucleotides do not form secondary structures that might hinder the cohesive ends for annealing to the handles, or that might compete with hairpin formation.
- Preferentially use the shorter cohesive end (i.e. Tsp45I) for the shorter oligonucleotide (i.e. the oligonucleotide not containing the loop).

D.5.3 *Annealing of final construct*

The annealing of the long handles to the oligonucleotides is performed in two steps. First each handle is incubated with its corresponding oligo using a high temperature protocol. Then the two reactions are mixed together and annealed using a lower temperature protocol[13]. The annealing protocol follows:

Annealing reaction:		
Handle A dig-tailed (~115 ng/µl)	2.8 µl (~0.6 pmols)	-
Handle B biotin-tailed (~210 ng/µl)	-	1 µl (~0.6 pmols)
oligo I with Tsp45I-end (e.g. EcoRIlong-I) (1 µM)	-	0.6 µl (~0.6 pmols)
oligo II with TspRI-end (e.g. EcoRIlong-IIB) (1 µM)	0.6 µl (~0.6 pmols)	-
1 M Tris	0.3 µl	0.3 µl
1 M NaCl	1 µl	1 µl
MilliQ water	5.3 µl	7.1 µl
Total volume:	10 µl	10 µl

13 Although the whole annealing protocol might be performed in a single step.

Three-way junction sequence
5'-*AATT***CTCA**GGGCACGCAGGCGGTTTGGAAACAAGCTTGAGTCTC
GTAGAGGGTTCGCCTGGACGAAGACTGACGCTCAGGTGCCC**ACACA***AGCT*-3'

Table 11: Sequence containing the three-way junction cloned between the EcoRI and HindIII sites of pBR322 plasmid. In italics the restriction enzyme sequences and in bold the ssDNA spacers for the pulling experiments.

Protocol: Incubate each reaction at 70°C for 1', decrease by 0.5°C every 20" down to 40°C. Hold at 4070°C, mix both reactions and decrease by 0.5°C every 20" down to 10°C.

Finally, the full molecular construct is ligated in an overnight reaction:

Ligation reaction:	
Annealing product[14]	20 µl
MilliQ water	20 µl
10X T4 DNA ligase buffer	5 µl
10 mM ATP	2 µl
T4 DNA ligase (400 U/µl)	2.5 µl
Total volume:	50 µl

Protocol: Overnight reaction at 16°C.

The ligated hairpin is then thermally inactivated (10' at 65°C) and aliquoted for storage at -20°C (1-2 µl EDTA 0.5 M can also be added).

D.6 THREE-WAY RNA JUNCTION WITH HETEROHANDLES

To perform pulling experiments of an RNA three-way junction, we cloned a minimal sequence (77 nucleotides + 9 ssDNA linkers) of the 16S ribosomal RNA that folds into this structure. Indeed, this region is the binding site of the ribosomal S15 protein that specifically recognizes and binds the three-way junction fold, playing a main role in the assembly of the small subunit of the prokaryotic ribosome. This sequence of interest (Table 11) was cloned between the EcoRI and HindIII sites of plasmid pBR322 (Figure 84**a**), and the flanking regions of the plasmid were used to engineer DNA/RNA heteroduplex handles for single-molecule pulling experiments (Figure 84**b**).

14 The ratio of volumes is relevant, as a high salt concentration decreases ligation efficiency.

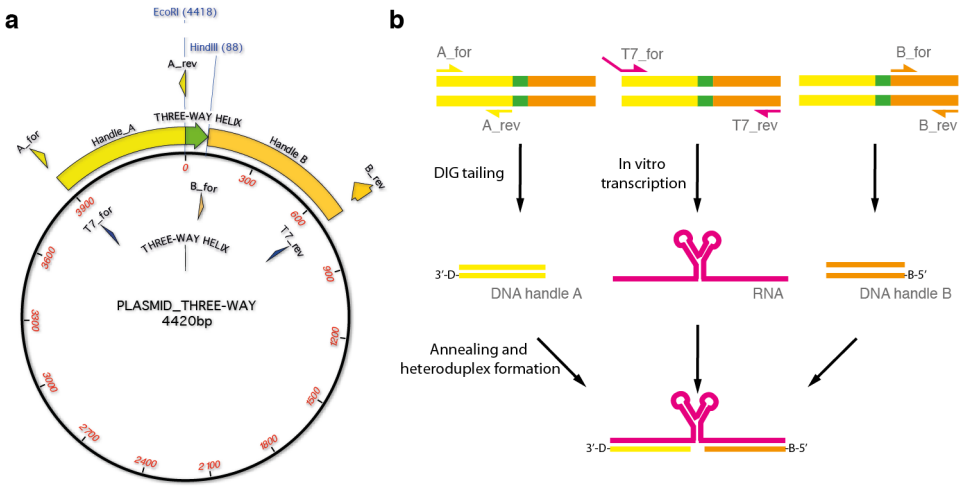

Figure 84: Scheme of RNA synthesis. (**a**) The sequence that contains the RNA three way junction (green) is cloned in pBR322 plasmid and the flanking regions are used to create the handles for pulling (yellow and orange). Arrows indicate the binding site of the primers. (**b**) Different regions of the plasmid are PCR amplified (top) to create the template for RNA transcription and the handles for tethering. The handles (yellow and orange) are differentially tagged with biotin and digoxigenin and annealed to the full RNA transcript (magenta) to create the final construct for manipulation with optical tweezers.

D.6.1 *Preparation of the recombinant plasmid*

To obtain the recombinant plasmid for RNA transcription of the molecular construct, the sequence of interest (Table 11) is cloned between the HindIII and EcoRI sites of pBR322 plasmid. To prepare the vector, pBR322 is incubated with the two restriction enzymes and with Calf Intestinal Alkaline Phosphatase (CIP):

Restriction and dephosphorylation of the vector:	
MilliQ water	131 µl
pBR322 vector (0.25 µg/µl)	40 µl
10X NEB EcoRI buffer	20 µl
CIP enzyme	4 µl
HindIII enzyme	2.5 µl
EcoRI enzyme	2.5 µl
Total volume:	200 µl
Incubate 1h at 37°C.	

The digested vector is then gel purified (0.8% agarose gel, 1XTBE) using Sybr Safe staining and blue illumination to minimize DNA damage and nicking. The vector is extracted using Qiaquick gel extraction kit (Qiagen). Next, we purchase two complementary DNA

oligonucleotides that self-assemble into the sequence of interest (Table 11) leaving an EcoRI and HindIII compatible overhang at each end (in italics in Table 11). The oligonucleotides are purchased 5'-phosphorylated and PAGE-purified. Alternatively they can be phosphorylated using T4 DNA kinase:

Phosphorylation and annealing of insert:	
MilliQ water	42 µl
oligo A (100 µM in Tris 10 mM)	20 µl
oligo B (100 µM in Tris 10 mM)	20 µl
T4 polynucleotide kinase (10 U/µl)	8 µl
10X T4 ligase buffer (not kinase buffer)	10 µl
Total volume:	100 µl

Incubate 1h at 37°C. Then anneal oligos using temperature gradient (95°C to 4°C in 2h)

Then the linearized and dephosphorylated vector is ligated with the dsDNA insert in an overnight reaction (generally a ~1:100 vector:insert ration is used). In this way, the recombinant plasmid contains the DNA insert that is the template of the desired RNA molecule.

Ligation of recombinant plasmid:	
MilliQ water	4.6 µl
DNA insert (1 µg/µl)	3 µl
digested vector (100 ng/µl)	0.4 µl
T4 DNA ligase (20 U/µl)	1 µl
10X T4 ligase buffer	1 µl
Total volume:	10 µl

Incubate overnight at 19°C (or at least 3h.)

D.6.2 *Transformation of competent cells with recombinant plasmid*

For this synthesis the plasmid was obtained from a genomic services provider (DNA2.0, Eurofins), eluted from a paper filter using 100 µl Tris-Cl 10 mM pH7.5 and quantified (~20 ng/µl). Ultra competent cells XL10-GOLD (Quickchange II XL site-directed mutagenesis kit) were transformed with the recombinant plasmid using the kit instructions but replacing the NZY+ Broth with LB. The transformation process was performed as follows (work in sterile conditions):

1. Aliquot 45 µl ultra competent cells on microtubes that have been pre-cooled on ice.

2. Add 2 µl β-mercaptoethanol. Mix and and incubate for 10′ on ice, mixing every 2′.

3. Add 1 µl plasmidic DNA (20 ng/µl). Mix well the transformation reaction and incubate 30′ on ice. Meanwhile, preheat LB at 42°C.

4. Heat-pulse the transformation reaction at 42°C for exactly 30″. Incubate on ice for 2′.

5. Add 1 ml LB preheated at 42°C and incubate on shaker at 37°C at 225-250 rpm for 1 hour.

6. Inoculate transformed cells into agar plates by (i) streak culturing using and inoculation loop, or (ii) by spreading increasing dilutions of the preparation with a glass culture spreader into separate agar plates. In this case we followed the second technique and prepared 3 different dilutions finding isolated colonies on the most diluted sample:
   - Centrifuge 400 µl preparation for 2′, remove 300 µl supernatant and inoculate the remaining 100 µl.
   - Inoculate 100 µl preparation.
   - Inoculate 10 µl preparation + 90 µl LB.

7. Each dilution is placed on a dry agar plate and spread with a glass spreader that is sterilized by immersion in 95% ethanol and burnt off with a Bunsen flame before and after inoculation (cool down by touching an agar plate before culturing). Allow liquid to dry and incubate overnight at 37°C with agar facing down and without agitation.

LB preparation for liquid culture:	
Triptone	5 grs
Yeast extract	2.5 gr
NaCl	5 grs
Total volume:	500 ml (MilliQ water)
Autoclave and use in sterile conditions.	

Petri dish preparation for cell culture:	
Triptone	2 grs
Yeast extract	1 gr
NaCl	2 grs
Agar	3 grs (for 1.5%)
Total volume:	200 ml (MilliQ water)

Protocol: Autoclave, let cool down to ~50°C and add ampicillin to 100 μg/ml (200μlof a 100 mg/ml stock) in sterile conditions. Apply ~15-20 ml per dish. Store solidified plates at 4°C facing down in a sealed bag.

D.6.3 *Selection of recombinants (inoculation of colonies in liquid culture and plasmid isolation)*

Place 3 ml sterile LB medium in a 15 ml Falcon tube and add ampicillin at at concentration of 50 μg/ml in sterile conditions (prepare 5 tubes). Inoculate a single colony in each tube, by picking it up from the recombinant agar plates with a sterile micropipette tip or tooth stick. Incubate the liquid cultures overnight at 37°C with agitation and with the cap slightly loose.

After growing the cultures overnight, perform a miniprep extraction using a commercially available kit (Qiaprep from Qiagen, GenElute from Sigma). In this case, plasmids were eluted in 100 μl Tris-Cl 5 mM, recovering ~30-40 ng/μl per sample. The transformation of the plasmid was initially checked by performing endonuclease restriction assays (SalI+PstI, HindIII+PstI) and then by plasmid sequencing from a genomic services provider (Servicios Genomicos).

Restriction assays:	
MilliQ water	6 μl
recombinant plasmid (~30 ng/μl)	10 μl
10X NEB buffer 2 *(or 10X buffer SureH)*	2.5 μl
10X BSA, 1:10 dilution	2.5 μl
SalI enzyme, 20 U/μl	2 μl
HindIII enzyme, 20 U/μl *(or PstI, 20 U/μl)*	2 μl
Total volume:	25 μl

Incubate 1h at 37°C.

Analyze in 1.2% agarose gel in TBE 1X using MW Marker VII from Roche. Expected bands are 2960/1460 bp (SalI+PstI), 3582/780/52 bp (PstI+HindIII).

The bacteria grown on the LB agar plate can be stored at 4°C for a few weeks. To keep a long term stock of transformed bacteria, a glycerol stock can be prepared and stored for years at -80°C. The glycerol stock is prepared by adding 500 μl of an overnight liquid culture of the transformed bacteria to 500 μl of 50% glycerol in a 2 ml microtube (mix well but gently). To recover bacteria from the glycerol stock, open the tube and take some of the frozen stock with a pipette tip without letting the glycerol thaw. Then grow bacteria in an LB agar plate.

D.6.4 *PCR amplification of DNA handles and template for in vitro transcription*

The recombinant vector is used both to create the template for *in vitro* transcription of the RNA structure, and to create the DNA handles for pulling on the molecular construct (Figure 84).

The template for *in vitro* transcription is created by PCR amplification of the three-way junction sequence (Figure 84**b**, green) together with the flanking regions that conform the handles for pulling experiments (Figure 84**b**, yellow, orange). The forward primer used for PCR amplification (Table 12) contains a 5′ non-hibridizing region designed to include the T7 promoter region in the amplicon.

PCR conditions for transcription template (1X):	
MilliQ water	30 µl
10X Optibuffer	10 µl
2 mM dNTPs	6 µl
50 mM $MgCl_2$	4 µl
5X HiSpec additive	20 µl
10 µM primer T7-for	10 µl
10 µM primer T7-rev	10 µl
1 ng/µl recombinant plasmid/PstI	5 µl
Ecotaq plus (40 U/µl), in Hot Start	5 µl
Total volume:	100 µl
PCR Procotol:	95°C for 1′; 30 x (95°C for 45″, 60°C for 1′, 72°C for 1′30″) ; 72°C for 5′.

A total volume of 6X PCR reactions is performed (scale as needed), purified using the QIAquick PCR purification kit and eluted in 50 µl Tris-Cl 10 mM, recovering ~174 ng/µl. The amplification product (1228 bp) is then gel purified (1.2% agarose gel, 1X TBE) using Sybr Safe staining and blue illumination to minimize DNA damage and nicking[15]. The DNA fragment is extracted from the gel using the QIAquick gel extraction kit (~90 ng/µl eluted in ~50 µl Tris-Cl 10 mM).

The DNA handles to pull on the RNA molecular construct are also obtained by PCR amplification of the recombinant plasmid (Figure 84). Primers for PCR extension of the handle regions (Table 12) are designed to create a handle that can hybridize with one end of the RNA transcript (handle A, 535bp, yellow in Figure 84), and a handle able to hybrize with the other end of the transcript (handle B, 599 bp, orange in Figure 84). The reverse primer for handle B contains a digoxigenin at its 5′ end for tethering of the molecular construct. After PCR amplification handle A is tailed with biotin at its 3′ ends

15 Alternatively, the fragment can be extracted by partially staining the gel with EtBr to localize and mark the band position, and then isolate the band from the unstained region.

Name	Primer sequence
A-for	5′-ggaattccGACTGGTGAGTACTCAACCAAGTC-3′
A-rev	5′-ATTCTTGAAGACGAAAGGGC-3′
B-for	5′-GCTTTAATGCGGTAGTTTATCACAG-3′
B-rev	5′-Digoxigenin-GCATTAGGAAGCAGCCCAGTAGTAGG-3′
T7-for	5′-taatacgactcactatagggaCTGGTGAGTACTCAACCAAGTC-3′
T7-rev	5′-TAGGAAGCAGCCCAGTAGTAGG-3′

Table 12: Primers used for the amplification of the handles and the transcription of the molecular construct. Lower case sequences correspond to non-hibridizing fragments (in T7-for to introduce the T7 promoter region, in A-for to introduce a ssDNA spacer for bead tethering -a similar spacer is found in handle B-).

following the protocol described below. Note that the primers for PCR amplification of the handles do not exactly overlap with those of the RNA template for transcription. This is done in order to leave 3-4 ssDNA bases between the digoxigenin/biotin tags and the DNA/RNA heterohandle (ssDNA spacers).

PCR conditions for DNA handles (1X):	
MilliQ water	54 µl
10X Thermopol buffer	10 µl
2 mM dNTPs	10 µl
10 µM primer A-for (B-for)	10 µl
10 µM primer A-rev (B-rev)	10 µl
1 ng/µl recombinant plasmid/PstI	5 µl
Vent DNA polymerase NEB (2 U/µl), in Hot Start	1 µl
Total volume:	100 µl
PCR Procotol:	95°C for 1′; 30 x (95°C for 45″, 56°C for 1′, 72°C for 1′30″); 72°C for 5′.

A total volume of 10X PCR reactions is performed (scale as needed) and purified using the Wizard SV gel and PCR clean up system (eluted in 50 µl Tris-Cl 10 mM; yields: Handle A ~750 ng/µl, Handle B ~1.1 µg/µl). The products are gel purified (1.2% agarose gel, 1X TBE) using Sybr Safe staining and blue illumination to minimize DNA damage and nicking. The DNA fragment is extracted from the gel using the QIAquick gel extraction kit (eluted in 50 µl RNase fee water; yields: Handle A ~420 ng/µl, Handle B ~500 µg/µl).

D.6.4.1 *Tailing of handle A with biotin*

Handle A is tailed with biotins at its 3′-ends in order to add a biotin-tag into the DNA strand that will conform the biotinylated DNA/RNA heterohandle for single molecule manipulation. The purified PCR product is tailed using the 3′-5′ exonuclease activity of T4 DNA poly-

merase:

Handle A tailing reaction:	
MilliQ water	34 µl
purified Handle A (~16 µg)	40 µl
T4 DNA polymerase buffer (NEB buffer 2)	10 µl
BSA 100X	1 µl
1 mM Biotin-16-dUTP	10 µl
T4 DNA polymerase enzyme	5 µl
Total volume:	100 µl
Incubate at room temperature for 20 min.	

The labeled handle is then purified using the SV gel and clean up system (or equivalent) and eluted in 50 µl RNase free water (~ 234 ng/mul).

D.6.5 *In vitro transcription of RNA using T7 RNA polymerase*

Transcription of the DNA template is performed using the T7 Megascript Kit (Ambion). The PCR amplified template is first concentrated by centrifugal evaporation or ethanol precipitation and eluted in 20 µl RNase free water (~ 222 ng/µl). Transcription is performed according to the kit instructions but with a higher concentration of DNA template:

In vitro transcription of PCR fragment:		
	reaction	control
Nuclease free water	5 µl	6 µl
Ribonucleotides mix	24 µl	8 µl
10X reaction buffer	6 µl	2 µl
pTri-Xef control DNA (0.5 mg/ml)	0 µl	2 µl
recombinant plasmid (~ 220 ng/µl)	19 µl	0 µl
T7 enzyme mix	6 µl	2 µl
Total volume:	60 µl	20 µl
Procotol: Keep reactives on ice except buffer. Prepare reaction at room temperature and incubate for 3 h at 37°C. Seal cap with parafilm. For short RNA transcripts (<300 nucleotides) perform reaction overnight.		

To stop the reaction and remove the template DNA, 3 µl TURBO DNase is added to the sample[16] and incubated for 15 min. To purify the RNA transcript lithium chloride precipitation is used, al-

16 (1 µl for the control reaction)

though for short structured DNA transcripts (<300 nucleotides) a phenol:chloroform extraction is preferred (see section D.8).

1. Add nuclease-free water (90 µl) and LiCL Precipitation solution (75 µl) to the reaction to precipitate the RNA. Mix and chill at -20°C for at least 1 hour.
2. Centrifuge at 4°C for 15′ at maximum speed.
3. Remove supernatant, wash the pellet with 1 ml 70% ethanol, and centrifuge.
4. Carefully remove the ethanol, let dry out and resuspend in 15 µl nuclease free water (yield: ~ 3 µg/µl). Store at -80°C (or -20°C).

Check RNA on a 4% denaturing polyacrylamide gel (Figure 85) or a 1.2% RNA denaturing agarose gel (TAE/formamide).

D.6.6 *Annealing of final construct (DNA/RNA heterohandle formation)*

Finally the RNA transcript is annealed to the DNA products to introduce the biotin and digoxigenin tags to each end of the molecular construct. This is done by using an annealing buffer in which DNA/RNA heteroduplex formation is thermodynamically more stable than B-helix DNA. The annealing buffer is prepared as follows:

Annealing buffer:	
Fomamide (Ambion)	800 µl
0.5 M EDTA pH8.0	2 µl
1 M Pipes pH6.3	40 µl
5 M NaCl	80 µl
Total volume:	922 µl µl

Addition of the handles and RNA transcript brings the concentration to the optimal annealing conditions[17].

Annealing reaction:	
Annealing buffer	80 µl
Handle A (~ 234 ng/µl)	9 (~ 6 pmols) µl
Handle B (~ 474 ng/µl)	5 (~ 6 pmols) µl
RNA transcript (~ 234 ng/µl)	1 (~ 8 pmols) µl
Total volume:	95 µl

Procotol: 85°C for 10′, 62°C for 1.5h, 52°C for 1.5h, ramp to 10°C in 10′.

17 Annealing conditions: 80% formamide, 1 mM EDTA pH8.0, 40 mM Pipes pH6.3, 0.4 M NaCl.

D.7 PREPARATION OF RNA GELS

- Native PAGE gel for DNA/RNA (for 5%):

30% Acrylamide/Bisacrylamide (1:29)	5 ml
TBE 10X	3 ml
MilliQ water	21.5 ml
10% APS	300 µl (double if old stock)
Total volume:	*30 ml.*

Swirl and add 30 µl TEMED for polymerization. Load samples using 10% Blue juice LB, run at 60V (6 V/cm). For non-denaturing gels, stain with EtBr or SybrSafe.

- Denaturing PAGE gel for RNA (for 6%):

30% Acrylamide/Bisacrilamide (1:19)	6 ml
TBE 10X	3 ml
Urea 7 M	12.6 grs
MilliQ water	as needed
Total volume:	*30 ml.*

Heat in microwave to melt urea, add 300 µl 10% APS and 30 µl TEMED for polymerization. Prerun gel for 15′ at 20 W. Mix samples with 10X Blue juice LB (60% formamide might be added), incubate 5′ at 90°C, and 5′ on ice. Run at 80 V (10 W). Stain with SybrGreen II (post-staining or added to the loading buffer).

- Band purification from PAGE gel using "soak and crush" method: Add 5 volumes TE for every volume of gel. Chop the gel with a pipette tip and incubate in a shaker at 37° C or RT for 4 h to overnight[18]. Centrifuge the sample and recover supernatant[19]. Acrylamide can be further removed by passing sample through a Whatman GF/C filter. Purify and concentrate the sample using a phenol:chloroform extraction (see section D.8)

- Denaturing agarose gel for RNA (Prepare 1X TAE agarose gel as usual an apply RNA sample as follows):

RNA sample + water	(e.g.) 3 µl
Formamide	6 µl ($\geqslant$ 60% V/V)
10X Loading buffer	1 µl
Total volume:	*10 µl*

Incubate sample for 5′ at 65°C, keep 5′ on ice and load into the gel. For denaturing gels, stain preferentially with SybrGreen II or EtBr.

18 A freeze-rapid thaw cycle (30′ at -80°C and 5′ at 90°C) helps increase yields due to acrylamide matrix breaking.

19 Further sample can be recovered by reincubating the acrylamide pellet with 0.5 volumes TE.

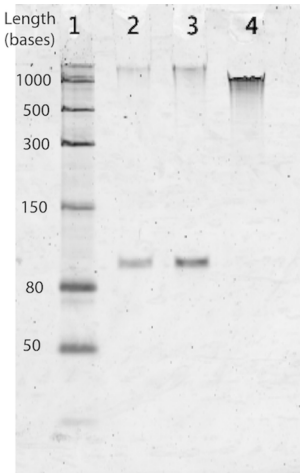

Figure 85: 6% denaturing PAGE of RNA transcripts. Lane 1 is a low range ssRNA ladder (NEB). Lanes 2,3 contain two concentrations of the short transcript of the three-way junction (94 nt), and lane 4 the long transcript of the three-way junction (with handles, ~1.2 kb).

D.8 PURIFICATION OF SHORT STRUCTURED RNA TRANSCRIPTS

To purify short RNA transcripts (<150 nt) with strong secondary structure it is not advisable to use LiCl precipitation or the Ambion MegaClear kit, as both methods are very efficient at purifying the sample from small structured RNA's. Consequently transcripts of the three-way junction without handles (94 bases with strong secondary structure) were purified using a phenol:chloroform extraction used for the purification of charged tRNA's. The protocol is as follows:

- Add MilliQ water to transcription reaction (20µl) up to 300 µl, and add 30 µl 3M KOAc pH5.3[20].
- Extract[21] with 300 µl cold acid phenol pH5.3 (4°C). Repeat this step at least once.
- Extract with 300µl chlorofolm, with 3 min centrifugation (optionally at 4°C too).
- Add 650 µl of 100% cold EtOH (stored at -20°C). Without incubation[22], directly centrifuge for 15 min (4°C).

20 prepare KOAc by dissolving potassium acetate in MilliQ water, and adjusting pH with glacial acetic acid, ~50% volume.

21 Mix well sample, centrifuge at top speed for 5 min (4°C), and extract aqueous phase (top layer). Cold acid phenol is most efficient at extracting RNA.

22 No incubation is needed for high concentration RNA, as incubation might cause salts coprecipitation. If no pellet is seen, the sample can be incubated at -20°C up to overnight.

- Wash with 1ml cold 70% EtOH (pipette up and down to wash). Air dry 5 to 10 min or spin vac for 5 min, and resuspend the pellet in 50 µl buffer or MilliQ water[23].
- Check RNA on a 6% denaturing polyacrylamide gel (Figure 85).

23 Or resuspend pellet in 50 µl 3 mM KOAc, and purify with G-25 column.

BIBLIOGRAPHY

[1] Abbondanzieri, E. A., Greenleaf, W. J., Shaevitz, J. W., Landick, R. & Block, S. M. Direct observation of base-pair stepping by RNA polymerase. *Nature* **438**, 460–465 (2005).

[2] Ritort, F. Single-molecule experiments in biological physics: methods and applications. *J. Phys.: Condens. Matter* **18**, R531 (2006).

[3] Collin, D. *et al.* Verification of the crooks fluctuation theorem and recovery of RNA folding free energies. *Nature* **437**, 231–234 (2005).

[4] Neuman, K. C. & Nagy, A. Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. *Nat. Methods* **5**, 491 (2008).

[5] Binnig, G., Quate, C. F. & Gerber, C. Atomic force microscope. *Phys. Rev. Lett.* **56**, 930 (1986).

[6] Moreno-Herrero, F., Colchero, J., Gomez-Herrero, J. & Baro, A. Atomic force microscopy contact, tapping, and jumping modes for imaging biological samples in liquids. *Phys. Rev. E* **69**, 031915 (2004).

[7] Bustamante, C., Rivetti, C. & Keller, D. J. Scanning force microscopy under aqueous solutions. *Curr. Opin. Struc. Biol.* **7**, 709–716 (1997).

[8] Kodera, N., Yamamoto, D., Ishikawa, R. & Ando, T. Video imaging of walking myosin V by high-speed atomic force microscopy. *Nature* **468**, 72–76 (2010).

[9] Ando, T., Uchihashi, T. & Kodera, N. High-speed AFM and applications to biomolecular systems. *Ann. Rev. Biophys.* **42**, 393–414 (2013).

[10] Rief, M., Gautel, M., Oesterhelt, F., Fernandez, J. M. & Gaub, H. E. Reversible unfolding of individual titin immunoglobulin domains by AFM. *Science* **276**, 1109–1112 (1997).

[11] Viani, M. B. *et al.* Small cantilevers for force spectroscopy of single molecules. *J. Appl. Phys.* **86**, 2258–2262 (1999).

[12] Bull, M. S., Sullan, R. M. A., Li, H. & Perkins, T. T. Improved single molecule force spectroscopy using micromachined cantilevers. *ACS nano* **8**, 4984–4995 (2014).

[13] Churnside, A. B. & Perkins, T. T. Ultrastable atomic force microscopy: Improved force and positional stability. *FEBS Lett.* (2014).

[14] Fernandez, J. M. & Li, H. Force-clamp spectroscopy monitors the folding trajectory of a single protein. *Science* **303**, 1674–1678 (2004).

[15] Garcia-Manyes, S., Liang, J., Szoszkiewicz, R., Kuo, T.-L. & Fernández, J. M. Force-activated reactivity switch in a bimolecular chemical reaction. *Nat. Chem.* **1**, 236–242 (2009).

[16] Grandbois, M., Beyer, M., Rief, M., Clausen-Schaumann, H. & Gaub, H. E. How strong is a covalent bond? *Science* **283**, 1727–1730 (1999).

[17] Moffitt, J. R., Chemla, Y. R., Izhaky, D. & Bustamante, C. Differential detection of dual traps improves the spatial resolution of optical tweezers. *Proc. Natl. Acad. Sci. USA* **103**, 9006–9011 (2006).

[18] Moffitt, J. R., Chemla, Y. R., Smith, S. B. & Bustamante, C. Recent advances in optical tweezers. *Biochemistry* **77**, 205 (2008).

[19] Smith, D. E. *et al.* The bacteriophage ϕ–29 portal motor can package DNA against a large internal force. *Nature* **413**, 748–752 (2001).

[20] Greenleaf, W. J., Woodside, M. T., Abbondanzieri, E. A. & Block, S. M. Passive all-optical force clamp for high-resolution laser trapping. *Phys. Rev. Lett.* **95**, 208102 (2005).

[21] Deufel, C., Forth, S., Simmons, C. R., Dejgosha, S. & Wang, M. D. Nanofabricated quartz cylinders for angular trapping: DNA supercoiling torque detection. *Nat. Methods* **4**, 223–225 (2007).

[22] Soltani, M. *et al.* Nanophotonic trapping for precise manipulation of biomolecular arrays. *Nat. Nanotechnol.* (2014).

[23] Lipfert, J., Hao, X. & Dekker, N. H. Quantitative modeling and optimization of magnetic tweezers. *Biophys. J.* **96**, 5040–5049 (2009).

[24] Neuman, K., Lionnet, T. & Allemand, J.-F. Single-molecule micromanipulation techniques. *Annu. Rev. Mater. Res.* **37**, 33–67 (2007).

[25] Yan, J., Skoko, D. & Marko, J. F. Near-field-magnetic-tweezer manipulation of single DNA molecules. *Phys. Rev. E* **70**, 011905 (2004).

[26] Fisher, J. *et al.* Thin-foil magnetic force system for high-numerical-aperture microscopy. *Rev. Sci. Instrum.* **77**, 023702 (2006).

[27] Manosas, M., Perumal, S. K., Croquette, V. & Benkovic, S. J. Direct observation of stalled fork restart via fork regression in the T4 replication system. *Science* **338**, 1217–1220 (2012).

[28] Manosas, M., Spiering, M. M., Zhuang, Z., Benkovic, S. J. & Croquette, V. Coupling DNA unwinding activity with primer synthesis in the bacteriophage T4 primosome. *Nat. Chem. Biol.* **5**, 904–912 (2009).

[29] Strick, T., Allemand, J. F., Croquette, V. & Bensimon, D. Twisting and stretching single DNA molecules. *Prog. Biophys. Mol. Bio.* **74**, 115–140 (2000).

[30] Strick, T. R., Croquette, V. & Bensimon, D. Single-molecule analysis of DNA uncoiling by a type ii topoisomerase. *Nature* **404**, 901–904 (2000).

[31] Koster, D. A., Palle, K., Bot, E. S., Bjornsti, M.-A. & Dekker, N. H. Antitumour drugs impede DNA uncoiling by topoisomerase I. *Nature* **448**, 213–217 (2007).

[32] Gosse, C. & Croquette, V. Magnetic tweezers: micromanipulation and force measurement at the molecular level. *Biophys. J.* **82**, 3314–3329 (2002).

[33] De Vlaminck, I. & Dekker, C. Recent advances in magnetic tweezers. *Ann. Rev. Biophys.* **41**, 453–472 (2012).

[34] Ribeck, N. & Saleh, O. A. Multiplexed single-molecule measurements with magnetic tweezers. *Rev. Sci. Instrum.* **79**, 094301 (2008).

[35] De Vlaminck, I. *et al.* Highly parallel magnetic tweezers by targeted DNA tethering. *Nano Lett.* **11**, 5489–5493 (2011).

[36] Ha, T. Single-molecule fluorescence resonance energy transfer. *Methods* **25**, 78–86 (2001).

[37] Yildiz, A. *et al.* Myosin V walks hand-over-hand: single fluorophore imaging with 1.5-nm localization. *Science* **300**, 2061–2065 (2003).

[38] Patterson, G., Davidson, M., Manley, S. & Lippincott-Schwartz, J. Superresolution imaging using single-molecule localization. *Annu. Rev. Phys. Chem.* **61**, 345 (2010).

[39] Huang, B., Bates, M. & Zhuang, X. Super resolution fluorescence microscopy. *Annu. Rev. Biochem.* **78**, 993 (2009).

[40] Hell, S. W. & Wichmann, J. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Opt. Lett.* **19**, 780–782 (1994).

[41] Gustafsson, M. G. Nonlinear structured-illumination microscopy: wide-field fluorescence imaging with theoretically unlimited resolution. *Proc. Natl. Acad. Sci. USA* **102**, 13081–13086 (2005).

[42] Rust, M. J., Bates, M. & Zhuang, X. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nat. Methods* **3**, 793–796 (2006).

[43] Betzig, E. *et al.* Imaging intracellular fluorescent proteins at nanometer resolution. *Science* **313**, 1642–1645 (2006).

[44] Roy, R., Hohng, S. & Ha, T. A practical guide to single-molecule FRET. *Nat. Methods* **5**, 507–516 (2008).

[45] Joo, C., Balci, H., Ishitsuka, Y., Buranachai, C. & Ha, T. Advances in single-molecule fluorescence methods for molecular biology. *Annu. Rev. Biochem.* **77**, 51–76 (2008).

[46] Comstock, M. J., Ha, T. & Chemla, Y. R. Ultrahigh-resolution optical trap with single-fluorophore sensitivity. *Nat. Methods* **8**, 335–340 (2011).

[47] Fu, H., Chen, H., Marko, J. F. & Yan, J. Two distinct overstretched DNA states. *Nucleic Acids Res.* gkq309 (2010).

[48] King, G. A. *et al.* Revealing the competition between peeled ssDNA, melting bubbles, and s-DNA during DNA overstretching using fluorescence microscopy. *Proc. Natl. Acad. Sci. USA* **110**, 3859–3864 (2013).

[49] Jackson, J. D. *Classical electrodynamics*, vol. 1 (Wiley-VCH, 1998), 3rd edn.

[50] Ashkin, A. *et al.* Acceleration and trapping of particles by radiation pressure. *Phys. Rev. Lett.* **24**, 156–159 (1970).

[51] Ashkin, A., Dziedzic, J., Bjorkholm, J. & Chu, S. Observation of a single-beam gradient force optical trap for dielectric particles. *Opt. Lett.* **11**, 288–290 (1986).

[52] Berns, M. W., Aist, J. R., Wright, W. H. & Liang, H. Optical trapping in animal and fungal cells using a tunable, near-infrared titanium-sapphire laser. *Exp. Cell Res.* **198**, 375–378 (1992).

[53] Barton, J., Alexander, D. & Schaub, S. Theoretical determination of net radiation force and torque for a spherical particle illuminated by a focused laser beam. *J. Appl. Phys.* **66**, 4594–4602 (1989).

[54] Wright, W. H., Sonek, G. & Berns, M. Parametric study of the forces on microspheres held by optical tweezers. *Appl. Optics* **33**, 1735–1748 (1994).

[55] Cheezum, M. K., Walker, W. F. & Guilford, W. H. Quantitative comparison of algorithms for tracking single fluorescent particles. *Biophys. J.* **81**, 2378–2388 (2001).

[56] Thompson, R. E., Larson, D. R. & Webb, W. W. Precise nanometer localization analysis for individual fluorescent probes. *Biophys. J.* **82**, 2775–2783 (2002).

[57] Rice, S. E., Purcell, T. J. & Spudich, J. A. Building and using optical traps to study properties of molecular motors. *Methods Enzymol.* **361**, 112–133 (2002).

[58] Gittes, F. & Schmidt, C. F. Interference model for back-focal-plane displacement detection in optical tweezers. *Opt. Lett.* **23**, 7–9 (1998).

[59] Neuman, K. C. & Block, S. M. Optical trapping. *Rev. Sci. Instrum.* **75**, 2787–2809 (2004).

[60] Smith, S. B., Cui, Y. & Bustamante, C. Optical-trap force transducer that operates by direct measurement of light momentum. *Methods Enzymol.* **361**, 134–162 (2002).

[61] Bustamante, C. J. & Smith, S. B. Light-force sensor and method for measuring axial optical-trap forces from changes in light momentum along an optic axis (2006). US Patent 7,133,132.

[62] Huguet, J. M. *et al.* Single-molecule derivation of salt dependent base-pair free energies in DNA. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 15431–15436 (2010).

[63] Lerman, L. Structural considerations in the interaction of DNA and acridines. *J. Mol. Biol.* **3**, 18–IN14 (1961).

[64] Chaires, J. B. A thermodynamic signature for drug–DNA binding mode. *Arch. Biochem. Biophys.* **453**, 26–31 (2006).

[65] Tse, W. C. & Boger, D. L. Sequence-selective DNA recognition: natural products and nature's lessons. *Chem. Biol.* **11**, 1607–1617 (2004).

[66] Waring, M. Variation of the supercoils in closed circular DNA by binding of antibiotics and drugs: evidence for molecular models involving intercalation. *J. Mol. Biol.* **54**, 247–279 (1970).

[67] Vladescu, I. D., McCauley, M. J., Nuñez, M. E., Rouzina, I. & Williams, M. C. Quantifying force-dependent and zero-force DNA intercalation by single-molecule stretching. *Nat. Methods* **4**, 517–522 (2007).

[68] Ren, J., Jenkins, T. C. & Chaires, J. B. Energetics of DNA intercalation reactions. *Biochemistry* **39**, 8439–8447 (2000).

[69] Haq, I. Thermodynamics of drug–DNA interactions. *Arch. Biochem. Biophys.* **403**, 1–15 (2002).

[70] Palchaudhuri, R. & Hergenrother, P. J. DNA as a target for anticancer compounds: methods to determine the mode of binding and the mechanism of action. *Curr. Opin. Biotech.* **18**, 497–503 (2007).

[71] Wang, A. H., Ughetto, G., Quigley, G. J. & Rich, A. Interactions between an anthracycline antibiotic and DNA: molecular structure of daunomycin complexed to d (CpGpTpApCpG) at 1.2–Å resolution. *Biochemistry* **26**, 1152–1163 (1987).

[72] Yakovchuk, P., Protozanova, E. & Frank-Kamenetskii, M. D. Base-stacking and base-pairing contributions into thermal stability of the DNA double helix. *Nucleic Acids Res.* **34**, 564–574 (2006).

[73] Frederick, C. A. *et al.* Structural comparison of anticancer drug–DNA complexes: adriamycin and daunomycin. *Biochemistry* **29**, 2538–2549 (1990).

[74] Lisgarten, J. N., Coll, M., Portugal, J., Wright, C. W. & Aymami, J. The antimalarial and cytotoxic drug cryptolepine intercalates into DNA at cytosine-cytosine sites. *Nat. Struct. Mol. Biol.* **9**, 57–60 (2001).

[75] Niyazi, H. *et al.* Crystal structures of λ-[ru (phen) 2dppz] 2+ with oligonucleotides containing TA/TA and AT/AT steps show two intercalation modes. *Nat. Chem.* **4**, 621–628 (2012).

[76] Leng, F., Chaires, J. B. & Waring, M. J. Energetics of echinomycin binding to DNA. *Nucleic Acids Res.* **31**, 6191–6197 (2003).

[77] Ughetto, G. *et al.* A comparison of the structure of echinomycin and triostin A complexed to a DNA fragment. *Nucleic Acids Res.* **13**, 2305–2323 (1985).

[78] Gao, Q. *et al.* Drug-induced DNA repair: X-ray structure of a DNA-ditercalinium complex. *Proc. Natl. Acad. Sci. USA* **88**, 2422–2426 (1991).

[79] Spielmann, H. P., Wemmer, D. E. & Jacobsen, J. P. Solution structure of a DNA complex with the fluorescent bis-intercalator TOTO determined by NMR spectroscopy. *Biochemistry* **34**, 8542–8553 (1995).

[80] Geierstanger, B. H. & Wemmer, D. E. Complexes of the minor groove of DNA. *Annu. Rev. Bioph. Biom.* **24**, 463–493 (1995).

[81] Van Dyke, M. M. & Dervan, P. B. Echinomycin binding sites on DNA. *Science* **225**, 1122–1127 (1984).

[82] Fox, K. R. & Waring, M. J. Stopped-flow kinetic studies on the interaction between echinomycin and DNA. *Biochemistry* **23**, 2627–2633 (1984).

[83] Bailly, C., Hamy, F. & Waring, M. J. Cooperativity in the binding of echinomycin to DNA fragments containing closely spaced CpG sites. *Biochemistry* **35**, 1150–1161 (1996).

[84] Glazer, A. N., Peck, K. & Mathies, R. A. A stable double-stranded DNA-ethidium homodimer complex: application to picogram fluorescence detection of DNA in agarose gels. *Proc. Natl. Acad. Sci. USA* **87**, 3851–3855 (1990).

[85] Glazer, A. N. & Rye, H. S. Stable dye-DNA intercalation complexes as reagents for high-sensitivity fluorescence detection. *Nature* **359**, 859–861 (1992).

[86] van Mameren, J. *et al.* Unraveling the structure of DNA during overstretching by using multicolor, single-molecule fluorescence imaging. *Proc. Natl. Acad. Sci. USA* **106**, 18231–18236 (2009).

[87] Günther, K., Mertig, M. & Seidel, R. Mechanical and structural properties of YOYO-1 complexed DNA. *Nucleic Acids Res.* **38**, 6526–6532 (2010).

[88] Tanious, F. A., Yen, S. F. & Wilson, W. D. Kinetic and equilibrium analysis of a threading intercalation mode: DNA sequence and ion effects. *Biochemistry* **30**, 1813–1819 (1991).

[89] Önfelt, B., Lincoln, P. & Nordén, B. A molecular staple for DNA: threading bis-intercalating [ru (phen) 2dppz] 2+ dimer. *J. Am. Chem. Soc.* **121**, 10846–10847 (1999).

[90] Wilhelmsson, L. M., Westerlund, F., Lincoln, P. & Nordén, B. DNA-binding of semirigid binuclear ruthenium complex δ, δ-[μ-(11, 11′-bidppz)(phen) 4ru2] 4+: Extremely slow intercalation kinetics. *J. Am. Chem. Soc.* **124**, 12092–12093 (2002).

[91] Nordell, P. & Lincoln, P. Mechanism of DNA threading intercalation of binuclear Ru complexes: uni-or bimolecular pathways depending on ligand structure and binding density. *J. Am. Chem. Soc.* **127**, 9670–9671 (2005).

[92] Howerton, B. S., Heidary, D. K. & Glazer, E. C. Strained ruthenium complexes are potent light-activated anticancer agents. *J. Am. Chem. Soc.* **134**, 8324–8327 (2012).

[93] Lee, J., Guelev, V., Sorey, S., Hoffman, D. W. & Iverson, B. L. NMR structural analysis of a modular threading tetraintercalator bound to DNA. *J. Am. Chem. Soc.* **126**, 14036–14042 (2004).

[94] Guelev, V., Sorey, S., Hoffman, D. W. & Iverson, B. L. Changing DNA grooves-a 1, 4, 5, 8-naphthalene tetracarboxylic diimide bis-intercalator with the linker (β-ala) 3-Lys in the minor groove. *J. Am. Chem. Soc.* **124**, 2864–2865 (2002).

[95] White, S., Szewczyk, J. W., Turner, J. M., Baird, E. E. & Dervan, P. B. Recognition of the four Watson–Crick base pairs in the DNA minor groove by synthetic ligands. *Nature* **391**, 468–471 (1998).

[96] Dervan, P. B. & Edelson, B. S. Recognition of the DNA minor groove by pyrrole-imidazole polyamides. *Curr. Opin. Struc. Biol.* **13**, 284–299 (2003).

[97] Holman, G. G., Zewail-Foote, M., Smith, A. R., Johnson, K. A. & Iverson, B. L. A sequence-specific threading tetra-intercalator with an extremely slow dissociation rate constant. *Nat. Chem.* **3**, 875–881 (2011).

[98] Bailly, C. & Chaires, J. B. Sequence-specific DNA minor groove binders. design and synthesis of netropsin and distamycin analogues. *Bioconjugate Chem.* **9**, 513–538 (1998).

[99] Neidle, S. DNA minor-groove recognition by small molecules. *Nat. Prod. Rep.* **18**, 291–309 (2001).

[100] Alberts, B. *et al.* *Molecular biology of the cell* (Garland Science, New York, 2002), 4th edn.

[101] Crow, S. D. *et al.* DNA sequence recognition by the antitumor drug ditercalinium. *Biochemistry* **41**, 8672–8682 (2002).

[102] Manning, G. S. The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides. *Q. Rev. Biophys.* **11**, 179–246 (1978).

[103] Record, M. T., Anderson, C. F. & Lohman, T. M. Thermodynamic analysis of ion effects on the binding and conformational equilibria of proteins and nucleic acids: the roles of ion association or release, screening, and ion effects on water activity. *Q. Rev. Biophys.* **11**, 103–178 (1978).

[104] Strekowski, L. & Wilson, B. Noncovalent interactions with DNA: an overview. *Mutat. Res.-Fund. Mol. M.* **623**, 3–13 (2007).

[105] Saccardo, P., Villaverde, A. & González-Montalbán, N. Peptide-mediated DNA condensation for non-viral gene therapy. *Biotechnol. Adv.* **27**, 432–438 (2009).

[106] Bloomfield, V. A. DNA condensation. *Curr. Opin. Struc. Biol.* **6**, 334–341 (1996).

[107] Braun, S. *et al.* Amyloid-associated nucleic acid hybridisation. *PLoS One* **6**, e19125 (2011).

[108] Komeda, S. *et al.* A third mode of DNA binding: phosphate clamps by a polynuclear platinum complex. *J. Am. Chem. Soc.* **128**, 16092–16103 (2006).

[109] Qu, X., Trent, J. O., Fokt, I., Priebe, W. & Chaires, J. B. Allosteric, chiral-selective drug binding to DNA. *Proc. Natl. Acad. Sci. USA* **97**, 12032–12037 (2000).

[110] Ohnmacht, S. A. *et al.* Discovery of new G-quadruplex binding chemotypes. *Chem. Commun.* **50**, 960–963 (2014).

[111] Neidle, S. Human telomeric G-quadruplex: The current status of telomeric G-quadruplexes as therapeutic targets in human cancer. *FEBS J.* **277**, 1118–1125 (2010).

[112] Hurley, L. H. DNA and its associated processes as targets for cancer therapy. *Nat. Rev. Cancer* **2**, 188–200 (2002).

[113] Balasubramanian, S., Hurley, L. H. & Neidle, S. Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? *Nat. Rev. Drug Discov.* **10**, 261–275 (2011).

[114] Pang, B. *et al.* Drug-induced histone eviction from open chromatin contributes to the chemotherapeutic effects of doxorubicin. *Nat. Commun.* **4**, 1908 (2013).

[115] Fink, A. L. Protein aggregation: folding aggregates, inclusion bodies and amyloid. *Fold. Des.* **3**, R9–23 (1998).

[116] Dima, R. & Thirumalai, D. Exploring protein aggregation and self-propagation using lattice models: Phase diagram and kinetics. *Protein Sci.* **11**, 1036–1049 (2002).

[117] Chiti, F. & Dobson, C. M. Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* **75**, 333–366 (2006).

[118] Tycko, R. Insights into the amyloid folding problem from solid-state NMR. *Biochemistry* **42**, 3151–3159 (2003).

[119] Gsponer, J. & Vendruscolo, M. Theoretical approaches to protein aggregation. *Protein Pept. Lett.* **13**, 287–293 (2006).

[120] Yoshimura, Y. *et al.* Distinguishing crystal-like amyloid fibrils and glass-like amorphous aggregates from their kinetics of formation. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 14446–14451 (2012).

[121] Dobson, C. M. Protein folding and misfolding. *Nature* **426**, 884–890 (2003).

[122] Eichner, T. & Radford, S. E. A diversity of assembly mechanisms of a generic amyloid fold. *Mol. Cell* **43**, 8–18 (2011).

[123] Chiti, F. *et al.* Designing conditions for in vitro formation of amyloid protofilaments and fibrils. *Proc. Natl. Acad. Sci. USA* **96**, 3590–3594 (1999).

[124] Chiti, F. *et al.* Kinetic partitioning of protein folding and aggregation. *Nat. Struct. Mol. Biol.* **9**, 137–143 (2002).

[125] Otzen, D. E., Kristensen, O. & Oliveberg, M. Designed protein tetramer zipped together with a hydrophobic alzheimer homology: a structural clue to amyloid assembly. *Proc. Natl. Acad. Sci. USA* **97**, 9907–9912 (2000).

[126] Schwartz, R. & King, J. Frequencies of hydrophobic and hydrophilic runs and alternations in proteins of known structure. *Protein Sci.* **15**, 102–112 (2006).

[127] Tartaglia, G. G., Pechmann, S., Dobson, C. M. & Vendruscolo, M. Life on the edge: a link between gene expression levels and aggregation rates of human proteins. *Trends Biochem. Sci.* **32**, 204 – 206 (2007).

[128] Chiti, F. *et al.* Studies of the aggregation of mutant proteins in vitro provide insights into the genetics of amyloid diseases. *Proc. Natl. Acad. Sci. USA* **99**, 16419–16426 (2002).

[129] Calamai, M. *et al.* Nature and significance of the interactions between amyloid fibrils and biological polyelectrolytes. *Biochemistry* **45**, 12806–12815 (2006).

[130] Di Domizio, J. *et al.* Binding with nucleic acids or glycosaminoglycans converts soluble protein oligomers to amyloid. *J. Biol. Chem.* **287**, 736–747 (2012).

[131] Campioni, S. *et al.* Salt anions promote the conversion of HypF-N into amyloid-like oligomers and modulate the structure of the oligomers and the monomeric precursor state. *J. Mol. Biol.* (2012).

[132] Motamedi-Shad, N. *et al.* Rapid oligomer formation of human muscle acylphosphatase induced by heparan sulfate. *Nat. Struct. Mol. Biol.* **19**, 547–554 (2012).

[133] Chiti, F., Stefani, M., Taddei, N., Ramponi, G. & Dobson, C. M. Rationalization of the effects of mutations on peptide andprotein aggregation rates. *Nature* **424**, 805–808 (2003).

[134] Kar, K., Jayaraman, M., Sahoo, B., Kodali, R. & Wetzel, R. Critical nucleus size for disease-related polyglutamine aggregation is repeat-length dependent. *Nat. Struct. Mol. Biol.* **18**, 328–336 (2011).

[135] Flory & J., P. *Principles of polymer chemistry* (Cornell University Press, 1953).

[136] Cohen, S. I. *et al.* Proliferation of amyloid-β42 aggregates occurs through a secondary nucleation mechanism. *Proc. Natl. Acad. Sci. USA* **110**, 9758–9763 (2013).

[137] Lambert, M. *et al.* Diffusible, nonfibrillar ligands derived from aβ1–42 are potent central nervous system neurotoxins. *Proc. Natl. Acad. Sci. USA* **95**, 6448–6453 (1998).

[138] Caughey, B. & Lansbury Jr, P. T. Protofibrils, pores, fibrils, and neurodegeneration: Separating the responsible protein aggregates from the innocent bystanders. *Annu. Rev. Neurosci.* **26**, 267–298 (2003).

[139] Ding, T. T., Lee, S.-J., Rochet, J.-C. & Lansbury, P. T. Annular α-synuclein protofibrils are produced when spherical protofibrils are incubated in solution or bound to brain-derived membranes. *Biochemistry* **41**, 10209–10217 (2002).

[140] Lansbury, P. T. Evolution of amyloid: what normal protein folding may tell us about fibrillogenesis and disease. *Proc. Natl. Acad. Sci. USA* **96**, 3342–3344 (1999).

[141] Jarrett, J. T., Berger, E. P. & Lansbury Jr, P. T. The carboxy terminus of the β–amyloid protein is critical for the seeding of amyloid formation: Implications for the pathogenesis of Alzheimer's disease. *Biochemistry* **32**, 4693–4697 (1993).

[142] Jarrett, J. T. & Lansbury Jr, P. T. Seeding "one-dimensional crystallization" of amyloid: a pathogenic mechanism in Alzheimer's disease and scrapie? *Cell* **73**, 1055–1058 (1993).

[143] Nilsberth, C. *et al.* 'arctic' APP mutation (E693G) causes Alzheimer's disease by enhanced Aβ protofibril formation. *Nat. Neurosci.* **4**, 887–893 (2001).

[144] Dahlgren, K. N. *et al.* Oligomeric and fibrillar species of amyloid-β peptides differentially affect neuronal viability. *J. Biol. Chem.* **277**, 32046–32053 (2002).

[145] Lashuel, H. A., Hartley, D., Petre, B. M., Walz, T. & Lansbury, P. T. Neurodegenerative disease: amyloid pores from pathogenic mutations. *Nature* **418**, 291–291 (2002).

[146] Wacker, J. L., Zareie, M. H., Fong, H., Sarikaya, M. & Muchowski, P. J. Hsp70 and Hsp40 attenuate formation of spherical and annular polyglutamine oligomers by partitioning monomer. *Nat. Struct. Mol. Biol.* **11**, 1215–1222 (2004).

[147] Campioni, S. *et al.* A causative link between the structure of aberrant protein oligomers and their toxicity. *Nat. Chem. Biol.* **6**, 140–147 (2010).

[148] Butterfield, S. M. & Lashuel, H. A. Amyloidogenic protein–membrane interactions: mechanistic insight from model systems. *Angew. Chem. Int. Ed. Engl.* **49**, 5628–5654 (2010).

[149] Puchalla, J., Krantz, K., Austin, R. & Rye, H. Burst analysis spectroscopy: A versatile single-particle approach for studying distributions of protein aggregates and fluorescent assemblies. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 14400–14405 (2008).

[150] Stranks, S. D. *et al.* Model for amorphous aggregation processes. *Phys. Rev. E* **80**, 051907 (2009).

[151] Nymeyer, H., García, A. E. & Onuchic, J. N. Folding funnels and frustration in off-lattice minimalist protein landscapes. *Proc. Natl. Acad. Sci. USA* **95**, 5921–5928 (1998).

[152] Morris, A. M., Watzky, M. A., Agar, J. N. & Finke, R. G. Fitting neurological protein aggregation kinetic data via a 2-step, minimal/"Ockham's razor" model: the Finke-Watzky mechanism of nucleation followed by autocatalytic surface growth. *Biochemistry* **47**, 2413–2427 (2008).

[153] Streets, A. M. & Quake, S. R. Ostwald ripening of clusters during protein crystallization. *Phys. Rev. Lett.* **104**, 178102 (2010).

[154] Borgia, M. B., Nickson, A. A., Clarke, J. & Hounslow, M. J. A mechanistic model for amorphous protein aggregation of immunoglobulin-like domains. *J. Am. Chem. Soc.* **135**, 6456–6464 (2013).

[155] Watzky, M. A. & Finke, R. G. Transition metal nanocluster formation kinetic and mechanistic studies. a new mechanism when hydrogen is the reductant: slow, continuous nucleation and fast autocatalytic surface growth. *J. Am. Chem. Soc.* **119**, 10382–10400 (1997).

[156] Coan, K. & Shoichet, B. Stoichiometry and physical chemistry of promiscuous aggregate-based inhibitors. *J. Am. Chem. Soc.* **130**, 9606–9612 (2008).

[157] Coan, K., Maltby, D., Burlingame, A. & Shoichet, B. Promiscuous aggregate-based inhibitors promote enzyme unfolding. *J. Med. Chem.* **52**, 2067–2075 (2009).

[158] Kauzmann, W. Of protein denaturation 1. *Adv. Protein Chem.* **14**, 1 (1959).

[159] Tanford, C. Contribution of hydrophobic interactions to the stability of the globular conformation of proteins. *J. Am. Chem. Soc.* **84**, 4240–4247 (1962).

[160] Tanford, C. The hydrophobic effect and the organization of living matter. *Science* **200**, 1012–1018 (1978).

[161] Tanford, C. How protein chemists learned about the hydrophobic factor. *Protein Sci.* **6**, 1358–1366 (1997).

[162] Dill, K. A. Dominant forces in protein folding. *Biochemistry* **29**, 7133–7155 (1990).

[163] Southall, N. T., Dill, K. A. & Haymet, A. A view of the hydrophobic effect. *J. Phys. Chem. B* **106**, 521–533 (2002).

[164] Helleday, T., Petermann, E., Lundin, C., Hodgson, B. & Sharma, R. A. DNA repair pathways as targets for cancer therapy. *Nat. Rev. Cancer* **8**, 193–204 (2008).

[165] Chenoweth, D. M. & Dervan, P. B. Allosteric modulation of DNA by small molecules. *Proc. Natl. Acad. Sci. USA* **106**, 13175–13179 (2009).

[166] Adhireksan, Z. *et al.* Ligand substitutions between ruthenium–cymene compounds can control protein versus DNA targeting and anticancer activity. *Nat. Commun.* **5** (2014).

[167] Warren, C. L. *et al.* Defining the sequence-recognition profile of DNA-binding molecules. *Proc. Natl. Acad. Sci. USA* **103**, 867–872 (2006).

[168] Fechter, E. J., Olenyuk, B. & Dervan, P. B. Design of a sequence-specific DNA bisintercalator. *Angew. Chem. Int. Ed. Engl.* **43**, 3591–3594 (2004).

[169] Westerlund, F., Wilhelmsson, L. M., Nordén, B. & Lincoln, P. Micelle-sequestered dissociation of cationic DNA-intercalated drugs: unexpected surfactant-induced rate enhancement. *J. Am. Chem. Soc.* **125**, 3773–3779 (2003).

[170] Phillips, D. R. & Crothers, D. M. Kinetics and sequence specificity of drug-DNA interactions: an in vitro transcription assay. *Biochemistry* **25**, 7355–7362 (1986).

[171] Lipfert, J., Klijnhout, S. & Dekker, N. H. Torsional sensing of small-molecule binding using magnetic tweezers. *Nucleic Acids Res.* **38**, 7122–7132 (2010).

[172] Paramanathan, T. *et al.* Mechanically manipulating the DNA threading intercalation rate. *J. Am. Chem. Soc.* **130**, 3752–3753 (2008).

[173] Kleimann, C. *et al.* Binding kinetics of bisintercalator Triostin A with optical tweezers force mechanics. *Biophys. J.* **97**, 2780–2784 (2009).

[174] Paramanathan, T., Vladescu, I., McCauley, M. J., Rouzina, I. & Williams, M. C. Force spectroscopy reveals the DNA structural dynamics that govern the slow binding of Actinomycin D. *Nucleic Acids Res.* **40**, 4925–4932 (2012).

[175] Paik, D. H. & Perkins, T. T. Dynamics and multiple stable binding modes of DNA intercalators revealed by single-molecule force spectroscopy. *Angew. Chem. Int. Ed. Engl.* **51**, 1811–1815 (2012).

[176] Almaqwashi, A. *et al.* Strong DNA deformation required for extremely slow DNA threading intercalation by a binuclear ruthenium complex. *Nucleic Acids Res.* **42** (2014).

[177] Sischka, A. *et al.* Molecular mechanisms and kinetics between DNA and DNA binding ligands. *Biophys. J.* **88**, 404–411 (2005).

[178] Murade, C., Subramaniam, V., Otto, C. & Bennink, M. L. Interaction of oxazole yellow dyes with DNA studied with hybrid optical tweezers and fluorescence microscopy. *Biophys. J.* **97**, 835–843 (2009).

[179] Hampshire, A. J., Rusling, D. A., Broughton-Head, V. J. & Fox, K. R. Footprinting: a method for determining the sequence selectivity, affinity and kinetics of DNA-binding ligands. *Methods* **42**, 128–140 (2007).

[180] Connaghan-Jones, K. D., Moody, A. D. & Bain, D. L. Quantitative DNase footprint titration: a tool for analyzing the energetics of protein–DNA interactions. *Nat. Protoc.* **3**, 900–914 (2008).

[181] Tulla-Puche, J. *et al.* Enzyme-labile protecting groups for the synthesis of natural products: Solid-phase synthesis of thiocoraline. *Angew. Chem. Int. Ed. Engl.* **125**, 5838–5842 (2013).

[182] Camunas-Soler, J. *et al.* Electrostatic binding and hydrophobic collapse of peptide–nucleic acid aggregates quantified using force spectroscopy. *ACS nano* **7**, 5102–5113 (2013).

[183] Albericio, F., Tulla-Puche, J., Auriemma, S. & Falciani, C. Orthogonal chemistry for the synthesis of thiocoraline-triostin hybrids. exploring their structure-activity relationship. *J. Med. Chem.* **56**, 5587–600 (2013).

[184] Zolova, O. E., Mady, A. S. & Garneau-Tsodikova, S. Recent developments in bisintercalator natural products. *Biopolymers* **93**, 777–790 (2010).

[185] Boger, D. L., Ichikawa, S., Tse, W. C., Hedrick, M. P. & Jin, Q. Total syntheses of thiocoraline and BE-22179 and assessment of their DNA binding and biological properties. *J. Am. Chem. Soc.* **123**, 561–568 (2001).

[186] Negri, A. *et al.* Antitumor activity, X-ray crystal structure, and DNA binding properties of thiocoraline A, a natural bisintercalating thiodepsipeptide. *J. Med. Chem.* **50**, 3322–3333 (2007).

[187] Shoichet, B. K. Screening in a spirit haunted world. *Drug Discov. Today* **11**, 607–615 (2006).

[188] Rocha, M., Ferreira, M. & Mesquita, O. Transition on the entropic elasticity of DNA induced by intercalating molecules. *J. Chem. Phys.* **127**, 105108 (2007).

[189] McGhee, J. D. & von Hippel, P. H. Theoretical aspects of DNA-protein interactions: co-operative and non-co-operative binding of large ligands to a one-dimensional homogeneous lattice. *J. Mol. Biol.* **86**, 469–489 (1974).

[190] Vladescu, I. D., McCauley, M. J., Rouzina, I. & Williams, M. C. Mapping the phase diagram of single DNA molecule force-induced melting in the presence of ethidium. *Phys. Rev. Lett.* **95**, 158102 (2005).

[191] Wakelin, L. & Waring, M. Kinetics of drug-DNA interaction: Dependence of the binding mechanism on structure of the ligand. *J. Mol. Biol.* **144**, 183–214 (1980).

[192] Reuter, M. & Dryden, D. T. The kinetics of YOYO-1 intercalation into single molecules of double-stranded DNA. *Biochem. Biophys. Res. Commun.* **403**, 225–229 (2010).

[193] Boger, D. L. & Saionz, K. W. DNA binding properties of key sandramycin analogues: systematic examination of the intercalation chromophore. *Bioorg. Med. Chem.* **7**, 315–321 (1999).

[194] Koch, S. J., Shundrovsky, A., Jantzen, B. C. & Wang, M. D. Probing protein-DNA interactions by unzipping a single DNA double helix. *Biophys. J.* **83**, 1098–1105 (2002).

[195] Hall, M. A. *et al.* High-resolution dynamic mapping of histone-DNA interactions in a nucleosome. *Nat. Struct. Mol. Biol.* **16**, 124–129 (2009).

[196] Huguet, J., Forns, N. & Ritort, F. Statistical properties of metastable intermediates in DNA unzipping. *Phys. Rev. Lett.* **103**, 248106 (2009).

[197] Bosco, A., Camunas-Soler, J. & Ritort, F. Elastic properties and secondary structure formation of single-stranded DNA at monovalent and divalent salt conditions. *Nucleic Acids Res.* **42**, 2064 (2014).

[198] Dawson, S., Malkinson, J. P., Paumier, D. & Searcey, M. Bisintercalator natural products with potential therapeutic applications: isolation, structure determination, synthetic and biological studies. *Nat. Prod. Rep.* **24**, 109–126 (2007).

[199] Galas, D. J. & Schmitz, A. DNAase footprinting a simple method for the detection of protein-DNA binding specificity. *Nucleic Acids Res.* **5**, 3157–3170 (1978).

[200] Berman, H. M. & Young, P. R. The interaction of intercalating drugs with nucleic acids. *Annu. Rev. Biophys. Bioeng.* **10**, 87–114 (1981).

[201] Erba, E. *et al.* Mode of action of thiocoraline, a natural marine compound with anti-tumour activity. *Br. J. Cancer.* **80**, 971 (1999).

[202] Esteller, M. Epigenetics in cancer. *N. Engl. J. Med.* **358**, 1148–1159 (2008).

[203] Chick, H. On the "heat coagulation" of proteins: Part IV. The conditions controlling the agglutination of proteins already acted upon by hot water. *J. Physiol.* **45**, 261–95 (1912).

[204] van Holde, K. E. Reflections on a century of protein chemistry. *Biophys. Chem.* **100**, 71–9 (2003).

[205] Masters, C. L. *et al.* Amyloid plaque core protein in Alzheimer disease and down syndrome. *Proc. Natl. Acad. Sci. USA* **82**, 4245–9 (1985).

[206] Ventura, S. & Villaverde, A. Protein quality in bacterial inclusion bodies. *Trends Biotechnol.* **24**, 179–185 (2006).

[207] Kopito, R. R. Aggresomes, inclusion bodies and protein aggregation. *Trends Cell Biol.* **10**, 524–530 (2000).

[208] Rosenberg, A. S. Effects of protein aggregates: an immunologic perspective. *AAPS J.* **8**, E501–E507 (2006).

[209] Yolamanova, M. *et al.* Peptide nanofibrils boost retroviral gene transfer and provide a rapid means for concentrating viruses. *Nat. Nanotechnol.* **8**, 130–136 (2013).

[210] den Engelsman, J. *et al.* Strategies for the assessment of protein aggregates in pharmaceutical biotech product development. *Pharm. Res.* 1–14 (2010).

[211] Aggeli, A. *et al.* Responsive gels formed by the spontaneous self-assembly of peptides into polymeric β-sheet tapes. *Nature* **386**, 259–62 (1997).

[212] Hamedi, M., Herland, A., Karlsson, R. H. & Inganäs, O. Electrochemical devices made from conducting nanowire networks self-assembled from amyloid fibrils and alkoxysulfonate pedot. *Nano Lett.* **8**, 1736–1740 (2008).

[213] Shaytan, A. K. *et al.* Self-assembling nanofibers from thiophene–peptide diblock oligomers: a combined experimental and computer simulations study. *ACS nano* **5**, 6894–6909 (2011).

[214] Carulla, N. *et al.* Molecular recycling within amyloid fibrils. *Nature* **436**, 554–558 (2005).

[215] Makin, O. S., Atkins, E., Sikorski, P., Johansson, J. & Serpell, L. C. Molecular basis for amyloid fibril formation and stability. *Proc. Natl. Acad. Sci. USA* **102**, 315–320 (2005).

[216] Petkova, A. T. *et al.* A structural model for Alzheimer's β-amyloid fibrils based on experimental constraints from solid state NMR. *Proc. Natl. Acad. Sci. USA* **99**, 16742–16747 (2002).

[217] Bucciantini, M. *et al.* Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature* **416**, 507–511 (2002).

[218] Macedo, B. *et al.* Nonspecific prion protein–nucleic acid interactions lead to different aggregates and cytotoxic species. *Biochemistry* **51**, 5402–5413 (2012).

[219] Cherny, D., Hoyer, W., Subramaniam, V. & Jovin, T. Double-stranded DNA stimulates the fibrillation of α-synuclein in vitro and is associated with the mature fibrils: an electron microscopy study. *J. Mol. Biol.* **344**, 929–938 (2004).

[220] Cohlberg, J., Li, J., Uversky, V. & Fink, A. Heparin and other glycosaminoglycans stimulate the formation of amyloid fibrils from α-synuclein in vitro. *Biochemistry* **41**, 1502–1511 (2002).

[221] Dale, T. Protein and nucleic acid together: a mechanism for the emergence of biological selection. *J. Theor. Biol.* **240**, 337–342 (2006).

[222] Feng, B. *et al.* Small-molecule aggregates inhibit amyloid polymerization. *Nat. Chem. Biol.* **4**, 197–199 (2008).

[223] Raman, B. *et al.* Critical balance of electrostatic and hydrophobic interactions is required for β2-microglobulin amyloid fibril growth and stability. *Biochemistry* **44**, 1288–1299 (2005).

[224] Donia, M. & Hamann, M. T. Marine natural products and their potential applications as anti-infective agents. *Lancet Infect. Dis.* **3**, 338–348 (2003).

[225] Hamann, M. T. & Scheuer, P. J. Kahalalide F: a bioactive depsipeptide from the sacoglossan mollusk Elysia rufescens and the green alga Bryopsis sp. *J. Am. Chem. Soc.* **115**, 5825–5826 (1993).

[226] Garcia-Rocha, M., Bonay, P. & Avila, J. The antitumoral compound Kahalalide F acts on cell lysosomes. *Cancer Lett.* **99**, 43–50 (1996).

[227] Suarez, Y. *et al.* Kahalalide F, a new marine-derived compound, induces oncosis in human prostate and breast cancer cells. *Mol. Cancer Ther.* **2**, 863–872 (2003).

[228] Sewell, J. M. *et al.* The mechanism of action of Kahalalide F: variable cell permeability in human hepatoma cell lines. *Eur. J. Cancer* **41**, 1637–1644 (2005).

[229] Lopez-Macia, A., Jimenez, J. C., Royo, M., Giralt, E. & Albericio, F. Synthesis and structure determination of Kahalalide F. *J. Am. Chem. Soc.* **123**, 11398–11401 (2001).

[230] Scott, A. M., Wolchok, J. D. & Old, L. J. Antibody therapy of cancer. *Nat. Rev. Cancer* **12**, 278–287 (2012).

[231] Weiner, L. M. Building better magic bullets—improving unconjugated monoclonal antibody therapy for cancer. *Nat. Rev. Cancer* **7**, 701–706 (2007).

[232] Lowe, S. W. & Lin, A. W. Apoptosis in cancer. *Carcinogenesis* **21**, 485–495 (2000).

[233] Cotter, T. G. Apoptosis and cancer: the genesis of a research field. *Nat. Rev. Cancer* **9**, 501–507 (2009).

[234] Jordan, M. A. & Wilson, L. Microtubules as a target for anticancer drugs. *Nat. Rev. Cancer* **4**, 253–265 (2004).

[235] Janmaat, M. L., Rodriguez, J. A., Jimeno, J., Kruyt, F. A. E. & Giaccone, G. Kahalalide F induces necrosis-like cell death that involves depletion of ErbB3 and inhibition of Akt signaling. *Mol. Pharm.* **68**, 502 (2005).

[236] Molina-Guijarro, J. *et al.* Irvalec inserts into the plasma membrane causing rapid loss of integrity and necrotic cell death in tumor cells. *PLoS One* **6**, e19042 (2011).

[237] Finkel, E. BIOMEDICINE: Does cancer therapy trigger cell suicide? *Science* **286**, 2256 (1999).

[238] Teixido, C., Marés, R., Aracil, M., y Cajal, S. R. & Hernandez-Losa, J. Epithelial-mesenchymal transition markers and HER3 expression are predictors of elisidepsin treatment response in breast and pancreatic cancer cell lines. *PloS one* **8**, e53645 (2013).

[239] Serova, M. *et al.* Predictive factors of sensitivity to elisidepsin, a novel kahalalide f-derived marine compound. *Mar. Drugs* **11**, 944–959 (2013).

[240] Kang, Y. & Massagué, J. Epithelial-mesenchymal transitions: twist in development and metastasis. *Cell* **118**, 277–279 (2004).

[241] Thiery, J. P. Epithelial–mesenchymal transitions in tumour progression. *Nat. Rev. Cancer* **2**, 442–454 (2002).

[242] Tebbutt, N., Pedersen, M. W. & Johns, T. G. Targeting the ERBB family in cancer: couples therapy. *Nat. Rev. Cancer* **13**, 663–673 (2013).

[243] Váradi, T. *et al.* ErbB protein modifications are secondary to severe cell membrane alterations induced by elisidepsin treatment. *Eur. J. Pharmacol.* **667**, 91–99 (2011).

[244] Dennis, J. W., Granovsky, M. & Warren, C. E. Glycoprotein glycosylation and cancer progression. *BBA-Gen. Subjects* **1473**, 21–34 (1999).

[245] Hakomori, S. Glycosylation defining cancer malignancy: new wine in an old bottle. *Proc. Natl. Acad. Sci. USA* **99**, 10231–10233 (2002).

[246] Dube, D. H. & Bertozzi, C. R. Glycans in cancer and inflammation potential for therapeutics and diagnostics. *Nat. Rev. Drug Discov.* **4**, 477–488 (2005).

[247] Lopez, P. H. & Schnaar, R. L. Gangliosides in cell recognition and membrane protein regulation. *Curr. Opin. Struc. Biol.* **19**, 549–557 (2009).

[248] Bouchiat, C. *et al.* Estimating the persistence length of a worm-like chain molecule from force-extension measurements. *Biophys. J.* **76**, 409–413 (1999).

[249] Smith, S., Cui, Y. & Bustamante, C. Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules. *Science* **271**, 795–799 (1996).

[250] Wang, M. D., Yin, H., Landick, R., Gelles, J. & Block, S. M. Stretching DNA with optical tweezers. *Biophys. J.* **72**, 1335–1346 (1997).

[251] Baumann, C. G., Smith, S. B., Bloomfield, V. A. & Bustamante, C. Ionic effects on the elasticity of single DNA molecules. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 6185–6190 (1997).

[252] Li, I. & Walker, G. Signature of hydrophobic hydration in a single polymer. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 16527–16532 (2011).

[253] Frisken, B. J. Revisiting the method of cumulants for the analysis of dynamic light-scattering data. *Appl. Opt.* **40**, 4087–4091 (2001).

[254] Rodriguez-Mias, R. A. *NMR in drug discovery. From screening to structure-based design of antitumoral agents*. Ph.D. thesis, Universitat de Barcelona (2006).

[255] Ritort, F., Mihardja, S., Smith, S. B. & Bustamante, C. Condensation transition in DNA-polyaminoamide dendrimer fibers studied using optical tweezers. *Phys. Rev. Lett.* **96**, 118301 (2006). cond-mat/0605737.

[256] Leger, J. F. *et al.* Structural transitions of a twisted and stretched DNA molecule. *Phys. Rev. Lett.* **83**, 1066–1069 (1999).

[257] Forns, N. *et al.* Improving signal/noise resolution in single-molecule experiments using molecular constructs with short handles. *Biophys. J.* **100**, 1765–1774 (2011).

[258] Podgornik, R. & Jönsson, B. Stretching of polyelectrolyte chains by oppositely charged aggregates. *Europhys. Lett.* **24**, 501–506 (1993).

[259] Brower-Toland, B. *et al.* Mechanical disruption of individual nucleosomes reveals a reversible multistage release of dna. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 1960–1965 (2002).

[260] Skoko, D., Yan, J., Johnson, R. & Marko, J. Low-force DNA condensation and discontinuous high-force decondensation reveal a loop-stabilizing function of the protein Fis. *Phys. Rev. Lett.* **95**, 208101 (2005).

[261] Van Noort, J., Verbrugge, S., Goosen, N., Dekker, C. & Dame, R. T. Dual architectural roles of HU: formation of flexible hinges and rigid filaments. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 6969–6974 (2004).

[262] Todd, B. & Rau, D. Interplay of ion binding and attraction in DNA condensed by multivalent cations. *Nucleic Acids Res.* **36**, 501–510 (2008).

[263] Hormeño, S. *et al.* Condensation prevails over B-A transition in the structure of DNA at low humidity. *Biophys. J.* **100**, 2006–2015 (2011).

[264] Ross, C. & Poirier, M. What is the role of protein aggregation in neurodegeneration? *Nat. Rev. Mol. Cell Biol.* **6**, 891–898 (2005).

[265] Ceci, P. *et al.* DNA condensation and self-aggregation of Escherichia coli Dps are coupled phenomena related to the properties of the N-terminus. *Nucleic Acids Res.* **32**, 5935–5944 (2004).

[266] Ceci, P., Mangiarotti, L., Rivetti, C. & Chiancone, E. The neutrophil-activating Dps protein of Helicobacter pylori, HP-NAP, adopts a mechanism different from Escherichia coli Dps to bind and condense DNA. *Nucleic Acids Res.* **35**, 2247–2256 (2007).

[267] Hou, X. *et al.* Cisplatin induces loop structures and condensation of single dna molecules. *Nucleic Acids Res.* **37**, 1400–1410 (2009).

[268] Dansithong, W. *et al.* Cytoplasmic CUG RNA foci are insufficient to elicit key DM1 features. *PLoS One* **3**, e3968 (2008).

[269] Dickson, A. M. & Wilusz, C. J. Repeat expansion diseases: when a good RNA turns bad. *Wiley Interdiscip. Rev.: RNA* **1**, 173–192 (2010).

[270] García-López, A., Llamusí, B., Orzáez, M., Pérez-Payá, E. & Artero, R. In vivo discovery of a peptide that prevents CUG–RNA hairpin formation and reverses RNA toxicity in myotonic dystrophy models. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 11866–11871 (2011).

[271] Bitan, G. *et al.* Amyloid β–protein (Aβ) assembly: Aβ40 and Aβ42 oligomerize through distinct pathways. *Proc. Natl. Acad. Sci. USA* **100**, 330–5 (2003).

[272] Ahn, B. W. *et al.* Detection of β-amyloid peptide aggregation using DNA electrophoresis. *Anal. Biochem.* **284**, 401–405 (2000).

[273] Boland, K., Behrens, M., Choi, D., Manias, K. & Perlmutter, D. H. The serpin-enzyme complex receptor recognizes soluble, nontoxic amyloid-β peptide but not aggregated, cytotoxic amyloid-β peptide. *J. Biol. Chem.* **271**, 18032–18044 (1996).

[274] Pike, C. J., Burdick, D., Walencewicz, A. J., Glabe, C. G. & Cotman, C. W. Neurodegeneration induced by beta-amyloid peptides in vitro: the role of peptide assembly state. *J. Neurosci.* **13**, 1676–1687 (1993).

[275] Bertozzi, C. R. & Kiessling, L. L. Chemical glycobiology. *Science* **291**, 2357–2364 (2001).

[276] Agard, N. J., Baskin, J. M., Prescher, J. A., Lo, A. & Bertozzi, C. R. A comparative study of bioorthogonal reactions with azides. *ACS Chem. Biol.* **1**, 644–648 (2006).

[277] Sletten, E. M. & Bertozzi, C. R. Bioorthogonal chemistry: fishing for selectivity in a sea of functionality. *Angew. Chem. Int. Ed. Engl.* **48**, 6974–6998 (2009).

[278] Lee, A. *et al.* Direct observation of dynamic mechanical regulation of DNA condensation by environmental stimuli. *Angew. Chem. Int. Ed. Engl.* **126**, 10807–10811 (2014).

[279] Storm, C. & Nelson, P. Theory of high-force DNA stretching and overstretching. *Phys. Rev. E* **67**, 051906 (2003).

[280] Bustamante, C., Marko, J., Siggia, E. & Smith, S. Entropic elasticity of lambda-phage DNA. *Science* **265**, 1599–1600 (1994).

[281] Marko, J. F. & Siggia, E. D. Stretching DNA. *Macromolecules* **28**, 8759–8770 (1995).

[282] Liang, X., Kuhn, H. & Frank-Kamenetskii, M. D. Monitoring single-stranded DNA secondary structure formation by determining the topological state of DNA catenanes. *Biophys. J.* **90**, 2877–2889 (2006).

[283] Mishra, G., Giri, D. & Kumar, S. Stretching of a single-stranded DNA: Evidence for structural transition. *Phys. Rev. E* **79**, 031930 (2009).

[284] Zhang, Y., Zhou, H. & Ou-Yang, Z.-C. Stretching single-stranded DNA: interplay of electrostatic, base-pairing, and base-pair stacking interactions. *Biophys. J.* **81**, 1133–1143 (2001).

[285] Lacy, M. J. & Voss, E. W. Direct adsorption of ssDNA to polystyrene for characterization of the DNA/anti-DNA interaction, and immunoassay for anti-DNA autoantibody in New Zealand White mice. *J. Immunol. Methods.* **116**, 87–98 (1989).

[286] Manohar, S. *et al.* Peeling single-stranded DNA from graphite surface to determine oligonucleotide binding energy by force spectroscopy. *Nano Lett.* **8**, 4365–4372 (2008).

[287] Petrovykh, D. Y., Kimura-Suda, H., Whitman, L. J. & Tarlov, M. J. Quantitative analysis and characterization of DNA immobilized on gold. *J. Am. Chem. Soc.* **125**, 5219–5226 (2003).

[288] Petrovykh, D. Y. *et al.* Nucleobase orientation and ordering in films of single-stranded DNA on gold. *J. Am. Chem. Soc.* **128**, 2–3 (2006).

[289] Montanari, A. & Mézard, M. Hairpin formation and elongation of biomolecules. *Phys. Rev. Lett.* **86**, 2178–2181 (2001).

[290] Dessinges, M.-N. *et al.* Stretching single stranded DNA, a model polyelectrolyte. *Phys. Rev. Lett.* **89**, 248102 (2002).

[291] Wolfe, A. R. & Meehan, T. The effect of sodium ion concentration on intrastrand base-pairing in single-stranded DNA. *Nucleic Acids Res.* **22**, 3147–3150 (1994).

[292] McIntosh, D., Ribeck, N. & Saleh, O. Detailed scaling analysis of low-force polyelectrolyte elasticity. *Phys. Rev. E* **80**, 041803 (2009).

[293] Saleh, O., McIntosh, D., Pincus, P. & Ribeck, N. Nonlinear low-force elasticity of single-stranded DNA molecules. *Phys. Rev. Lett.* **102**, 068301 (2009).

[294] Rief, M., Clausen-Schaumann, H. & Gaub, H. E. Sequence-dependent mechanics of single DNA molecules. *Nat. Struct. Mol. Biol.* **6**, 346–349 (1999).

[295] Clausen-Schaumann, H., Rief, M., Tolksdorf, C. & Gaub, H. E. Mechanical stability of single DNA molecules. *Biophys. J.* **78**, 1997–2007 (2000).

[296] Danilowicz, C., Lee, C., Coljee, V. & Prentiss, M. Effects of temperature on the mechanical properties of single stranded DNA. *Phys. Rev. E* **75**, 030902 (2007).

[297] Odijk, T. Polyelectrolytes near the rod limit. *J. Polym. Sci. Pol. Phys.* **15**, 477–483 (1977).

[298] Skolnick, J. & Fixman, M. Electrostatic persistence length of a wormlike polyelectrolyte. *Macromolecules* **10**, 944–948 (1977).

[299] Manning, G. S. Counterion condensation on a helical charge lattice. *Macromolecules* **34**, 4650–4655 (2001).

[300] Caliskan, G. *et al.* Persistence length changes dramatically as RNA folds. *Phys. Rev. Lett.* **95**, 268303 (2005).

[301] Toan, N. M. & Thirumalai, D. Theory of biopolymer stretching at high forces. *Macromolecules* **43**, 4394–4400 (2010).

[302] McIntosh, D. & Saleh, O. Salt species-dependent electrostatic effects on ssDNA elasticity. *Macromolecules* **44**, 2328–2333 (2011).

[303] Shklovskii, B. I. Screening of a macroion by multivalent ions: correlation-induced inversion of charge. *Phys. Rev. E* **60**, 5802 (1999).

[304] Heilman-Miller, S. L., Thirumalai, D. & Woodson, S. A. Role of counterion condensation in folding of the *Tetrahymena* ribozyme. I. equilibrium stabilization by cations. *J. Mol. Biol.* **306**, 1157–1166 (2001).

[305] Bizarro, C., Alemany, A. & Ritort, F. Non-specific binding of Na+ and Mg2+ to RNA determined by force spectroscopy methods. *Nucleic Acids Res.* **40**, 6922–6935 (2012).

[306] Chen, H. *et al.* Ionic strength-dependent persistence lengths of single-stranded RNA and DNA. *Proc. Natl. Acad. Sci. USA* **109**, 799–804 (2012).

[307] Seol, Y., Skinner, G. M. & Visscher, K. Elastic properties of a single-stranded charged homopolymeric ribonucleotide. *Phys. Rev. Lett.* **93**, 118102 (2004).

[308] Saenger, W. *Principles of nucleic acid structure*, vol. 7 (Springer-Verlag New York, 1984).

[309] Bloomfield, V. A., Crothers, D. M. & Tinoco Jr, I. *Nucleic acids: structures, properties, and functions* (University science books, 2000).

[310] Manosas, M., Junier, I. & Ritort, F. Force-induced misfolding in RNA. *Phys. Rev. E* **78**, 061925 (2008).

[311] Borchardt, R., Kerns, E., Lipinski, C., Thakker, D. & Wang, B. *Pharmaceutical profiling in drug discovery for lead selection*, vol. 1 (American Assoc. of Pharm. Scientists, 2005).

[312] Kozikowski, B. A. *et al.* The effect of freeze/thaw cycles on the stability of compounds in DMSO. *J. Biomol. Screening* **8**, 210–215 (2003).

[313] McDonald, G. R. *et al.* Bioactive contaminants leach from disposable laboratory plasticware. *Science* **322**, 917 (2008).

[314] Nuijen, B. *et al.* Compatibility and stability of the investigational polypeptide marine anticancer agent Kahalalide F in infusion devices. *Invest. New Drugs* **19**, 273–281 (2001).

LIST OF PUBLICATIONS

- Camunas-Soler J, Frutos S, Bizarro CV, de Lorenzo S, Fuentes-Perez ME, Ramsch R, Vilchez S, Solans C, Moreno-Herrero F, Albericio F, Eritja R, Giralt E, Dev SB, Ritort F. Electrostatic binding and hydrophobic collapse of peptide-nucleic acid aggregates quantified using force spectroscopy. *ACS Nano*, 7(6):5102-13, 2013.
- Bosco A, Camunas-Soler J, Ritort F. Elastic properties and secondary structure formation of single-stranded DNA at monovalent and divalent salt conditions. *Nucleic Acids Res*, 42(3):2064-74, 2014.
- E. Dieterich*, Camunas-Soler J*, Ribezzi-Crivellari M, Seifert U, Ritort F. Single molecule measurement of the effective temperature in nonequilibrium steady states. *Submitted*.
- Camunas-Soler J, Manosas M, Tulla-Puche J, Albericio F, Ritort F. Single-molecule kinetics and footprinting of DNA bis-intercalation: the paradigmatic case of Thiocoraline. *Submitted*.
- Camunas-Soler J, Alemany A, Ritort F. Free energy measurements of ligands binding nucleic acids using fluctuation theorems. *In preparation*.

* - The authors equally contributed to the research.

ACRONYMS

AFM - Atomic Force Microscopy

APP - Amyloid Precursor Protein

APTES - 3-Aminopropyltriethoxysilane

CCD - Charge-coupled Device

CD - Circular Dichroism

CIP - Calf Intestinal Alkaline Phosphatase

DLS - Dynamic Light Scattering

DMSO - Dimethyl Sulfoxide

dsDNA - double stranded DNA

EDTA - Ethylenediaminetetraacetic Acid

EB - Electrostatic Binding

FDC - Force-Distance Curve

FEC - Force-Extension Curve

FJC - Freely Jointed Chain

FFT - Fast Fourier Transform

FRET - Fluorescence Resonance Energy Transfer

HB - Hydrophobic Collapse

KF - Kahalalide F

MGVH - McGhee Von Hippel

mPEG - methyl Polyethylene Glycol

NA - Numerical Aperture

NMR - Nuclear Magnetic Resonance

PBS - Phosphate Buffered Saline

PCS - Photon Correlation Spectrometer

PCR - Polymerase Chain Reaction

PSD - Position Sensing Detector

QPD - Quadrant Photodiode

SNR - Signal-to-Noise Ratio

ssDNA - single stranded DNA

SSB protein - single-stranded DNA binding protein

TBE - Tris Borate EDTA

TE - Tris-EDTA

TEM - Transmission Electron Microscopy

TIRF - Total Internal Reflection Fluorescence

WLC - Worm-Like Chain